

## Cloning Scheme for Generating pCMV-NS35

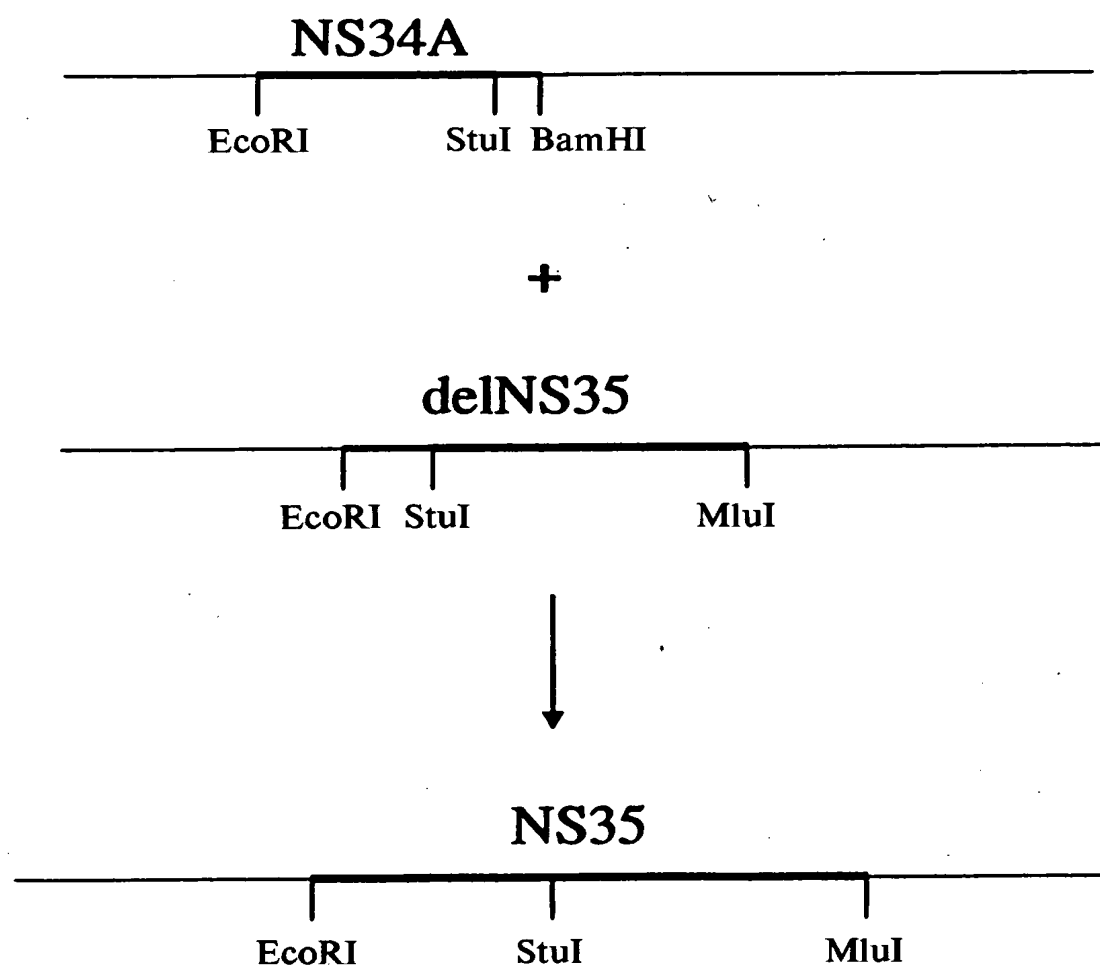


FIG. 1

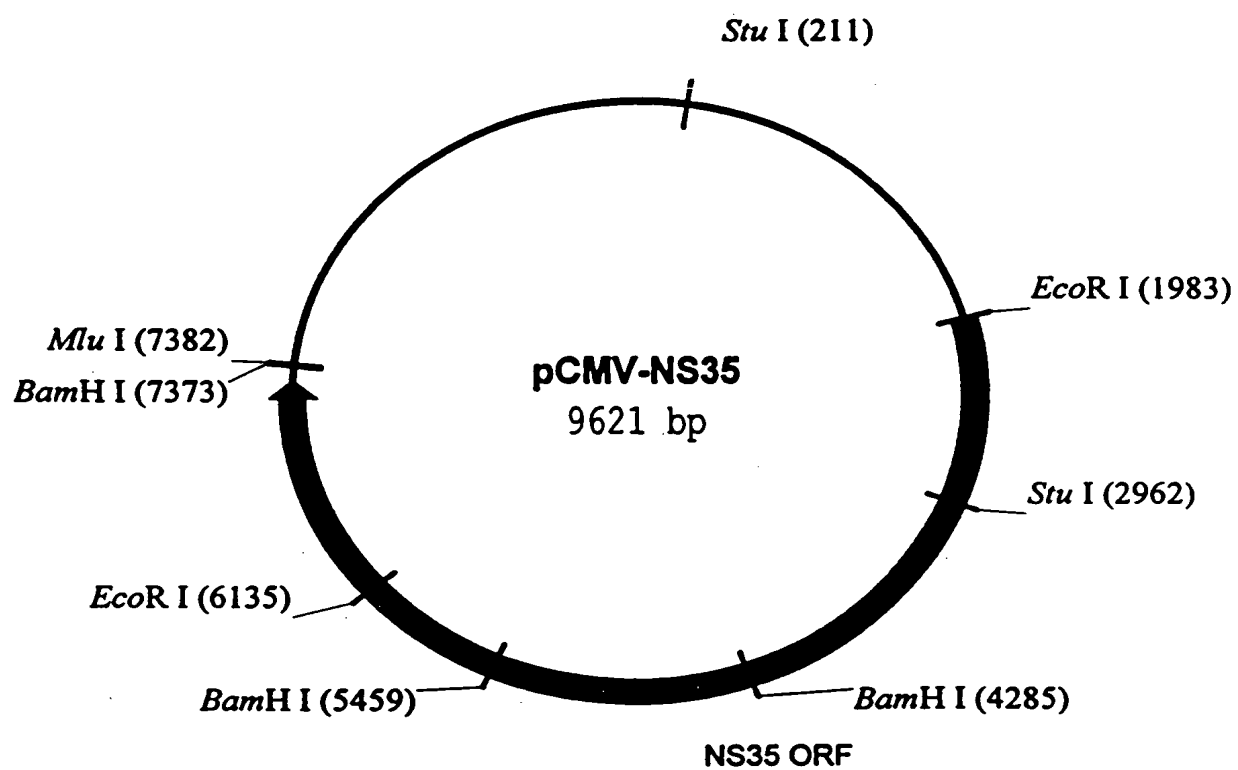


FIG. 2

pCMV-NS35

1	TCGGCGGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC	CTCTGCCAGT	GTGGAACAGA	CATTCCGCTA
81	GC	GGGGAGCA	GACAAGCCCG	TCAGGGCGCG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATG	CGGCATCAGA
	CGGCCCTCGT	CTGTTGGGGC	AGTCCCGCGC	AGTCGCCAC	AACGCCCCAC	AGCCCGGACC	GAATTGATAC	GGCGTAGTCT
161	GC	GAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTTCGA	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA
	CGTCTAACAT	GACTCTCAGG	TGGTATACTT	CGAAAAACGT	TTTCGGATCC	GGAGTTT	TCGGAGGAGT	GATGAAGACC
241	AATAGCTCAG	AGCCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAT	TAGTCAGCCA	TGGGGCGGAG	AATGGGCGGA
	TTATCGAGTC	TCCGGCTCCG	CCGGAGCCCG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT	ACCCCGCCTC	TTACCCGCCT
321	ACTGGGCGGG	GAGGGAATTA	TTGGCTATTG	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	TGACCCGCCC	CTCCCTTAAT	AACCGATAAC	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	TAGTAATCAA	TTACGGGGTC	ATTAGTTTCAT
	GTACAGGTTA	TACTGGCGGT	ACAACTGTAA	CTAATAACTG	ATCAATAAAT	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA
481	AGCCCATATA	TGGAGTTCCG	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	CCCGCCCATTT
	TCGGGTATAT	ACCTCAAGGC	GCAATGTATT	GAATGCCATT	TACCGGCGCG	ACCGACTGSC	GGGTTGCTGG	GGCGGGGTAA
561	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT
	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT	CCCTGAAAGG	TAACTGCAGT	TACCCACCTC	ATAAATGCCA
641	AAACTGCCCC	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCCC	CCTATTGCAG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTTCAT	GTAGTTTACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG

StuI  
-----

FIG. 3-Page 1

pCMV-NS35

```

721 GCCTGGCATT ATGCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
    CGGACCGTAA TACGGGTCAT GTACTGGAAT GGCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTGTGG AGTACACCAA TGGGGGTGGA TAGCGGTTTG ACTCAGGGG ATTTCCAAGT CTCCACCCCA
    GTACCACTAC GCCAAAACCG TCATGTGTT ACCCGCACCT ATGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGTGGGGT

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCCGTTGACG
    AACTGCAGTT ACCCTCAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACACCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
    GTTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGGGGCGG GAAACGGTGC ATTGGAACGC
    GGTAGGTGG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGG AGGCGCCGGC CCTTGCCAGG TAACCTTGCG

1121 GGATTCGCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
    CCTAAGGGCG ACGGTTCTCA CTGCATTTCAT GCGGATATC TGAGATATCC GTGTGGGGA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
    GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATT TTAGCCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACCTAT
    AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTTAT
    GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA

```

FIG. 3-Page 2

pCMV-NS35

```

1441 TATTACAA TTACATATA CAACAAGCC GTCCCCCGTG CCCGAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
    ATAAATGTTT AAGTGATAT GTTGTGGCG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCT AGAGGCTGTA

1521 CTCGGGTAGG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
    GAGCCCATGC ACAAGGCTG TACCCGAGAA GAGGCCATCG CCGCTCGAA GGTGAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGTCATG GTGGTCCGC AGTCCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
    CGCCGAGTAC CAGCGAGCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

1681 AGTGTGCCG ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
    TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
    CCTTCTGAAT TCCGTGCGG TCTTCTTCTA CGTCCGTGCA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
    ACGCCAGGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

+2 M A A
    EcoRI
    -----
1921 GACAGACTAA CAGACTGTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTACC ATGGCTGCAT
    CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACGGGCT TTGGTGCTTA CATGTCCAAG
    TAGGTGAGT CCCGATATC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACACGAAT GTACAGGTTT

```

pCMV-NS35

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G  
2081 GCTCATGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGCTACT CCACCTACGG  
CGAGTACCCT AGCTAGGATT GTAGTCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A  
2161 CAAGTTCCTT GCCAGCGGG GGTGCTCGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA  
GTTCAAGGAA CGGCTGCCG CCACGAGCCC CCGCGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T  
2241 CATCCATCTT GGGCATTGGC ACTGCTCTG ACCAAGCAGA GACTGCGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC  
GTAGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G  
2321 CTTCCGGCT CCGTCACTGT GCCCATCC AACATCGAGG AGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG  
GGAGGCCGA GGCAGTGACA CGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L  
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAA GAAGAAGTGC GACGAACCTCG  
GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA GAGTAGAGA CAGTAAGTTT CTCTTTCAG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G  
2481 CCGCAAAGCT GGTGCAATG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC  
GGCGTTTCA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTCGCCG

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C  
2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGCGGA CTTGACTCG GTGATAGACT GCAATACGTG  
CTACAACAGC AGCACCCTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

FIG. 3-Page 4

pCMV-NS35

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S  
2641 TCTACCCAG ACAGTCGATT TCAGCCTTGA CCTACCTTC ACCATTGAGA CAATCAGCT CCCCAAGAT GCTGTCTCCC  
ACAGTGGGTC TGTCAGCTAA AGTCGGAAG GGGATGGAAG TGTTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G  
2721 GCACTCAACG TCGGGCAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGCG  
CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AACACCCGTG GCCCCCTCGC GGGGAGGCGG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V  
2801 ATGTTCCGACT CGTCCGTCTT CTGTAGTGC TATGAGCGAG GCTGTGCTTG GTATGAGCTC ACGCCCGCGG AGACTACAGT  
TACAAAGCTGA GCAGGCAGGA GACACTCAGG ATACTGGGTC CGACACGAAC CATACTCGAG TCGGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T  
StuI

2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG  
ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCCTGG TAGAACTTAA AACCCCTCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q  
StuI

2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT ACCGTACCAA  
CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGGTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T  
3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCTT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCAC  
CGGTGGCACA CGCGATCCCG AGTTGGGGA GGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTGGGCTG

FIG. 3-Page 5

pCMV-NS35

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K  
3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCCCT GACGCACCCA GTACACCAAT  
GGAGGTACCC GGTGTGGGG ACGATATGTC TGAACCGCGA CAAGTCTTAC TTAGTGGGA CTGCGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L  
3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGTCTGTAC GAGCACCTGG GTGCTCGTTG GCGGCGTCTT GGCTGCTTTG  
TGTAAGTACTG TACGTACAGC CCGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC GCGCGCAGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D  
3281 GCGGGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGTCGT CTTGTCCGG AAGCCGGCAA TCATACCTGA  
CGCGGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCACGA GAACAGGCCC TTCGGCCGTT AGTATGGAAT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M  
3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC  
GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V  
3441 TCGCGGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC  
AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G  
3521 CAGACCAACT GGC AAAACT CGAGACCTTC TGGCGGAAGC ATATGTGGAA CTTTCATCAGT GGGATACAAT ACTTGGCGGG  
GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T  
3601 CTTGTCAAG CTGCCTGGTA ACCCGGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCGAGCCA CTAACCACTA  
GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTGGGT GATTGGTGAT



pCMV-NS35

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V  
3681 GCCAAACCCT CCTCTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC GCGGCCCG GTCCGCTAC TGCCTTGTG  
CGGTTGGGA GGAGAAGTTG TATAACCCG CCAACCACG ACGGTGAG CCGGGGGG CACGGCGATG ACGAAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A  
3761 GCGCTGGCT TAGCTGGCG CGCCATCGG AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCG  
CCGGACCGA ATCGACCGG GCGTAGCG TCACAACCTG ACCCTTCCA GGATATCTG TAGGAACGTC CCATACCGG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L  
3841 GCGCTGGCG GGAGCTCTTG TGCATTCAA GATCATGAG GGTGAGTCC CCTCCACGGA GGACCTGGT AATCTACTGC  
CCCGACCGC CCTCGAGAAC ACCGTAAGT CTAGTACTG CCACTCCAG GGAGTGCCT CCTGGACCAG TTAGATGAG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E  
3921 CCGCATCCT CTCGCCCGA GCGCTGATG TCGGCTGTG CTGTGCAGCA ATACTGCGC GGCACGTTGG CCGGGCGAG  
GGCGGTAGGA GAGCGGCCT CGGAGCATC AGCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E  
4001 GGGCAGTGC AGTGGATGAA CCGCTGATA GCGTGCCT CCGGGGGA CCATGTTCC CCCACGCACT ACGTCCCGA  
CCCCGTCAG TCACCTACTT GCGGCACTAT CGGAAGCGGA GGGCCCCCTT GTACAAAGG GGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W  
4081 GAGCGATGCA GTCGCCCGG TCACTGCCAT ACTCAGCAGC CTCAGTGTA CCCAGTCTCT GAGGCGACTG CACCACTGGA  
CTCGCTACGT CGACGGGCG AGTGACGGTA TGAGTCGTG GAGTGACATT GGTGCGAGGA CTCGCTGAC GTGGTCACT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D  
4161 TAAGCTCGA GTGTACCACT CCATGCTCG GTTCTGGCT AAGGACATC TGGGACTGGA TATGCGAGGT GTTGACCGAC  
ATTCGAGCCT CACATGGTGA GGTACGAGG CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G  
BamHI

4321 +2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
GGTCTGGCGA GGGCAGGCA TCATGCACAC TCGTGGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA  
CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCAGCGGTG ACACCTCGAC TCTAGTCACC TGTACAGTTT TTGCCCTGCT

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
4481 ACCCCCCTC CTGCGCCGAA CTACACCTTC GCGTATGGA GGTGTCTGC ACAGGAATAC GTGGAGATAA GGCAGGTGGG  
TGGGGGGAAG GACGGGGCTTT GATGTGGAAG CGCGATACCT CCCACACAGG TCTCCTTATG CACCTCTATT CCGTCCACCC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641 AATTGGACGG GGTGGCCTA CATAGGTTG CGGCCCTG CAAGCCTTG CTGCGGGAGG AGTATCATT CAGAGTAGGA  
 TTAACCTGCC CCACGGGAT GTATCCAAAC GCGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
4721 CTCCACGAAT ACCGGTAGG GTCCGAATTA CCTTGCAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA  
GAGTGCTTGA TGGGCGATCC CAGCGTTAAT GGAACGGTTCG GCGTTGGCCT GCACCGGAC AACTGCAGGT ACCAGTGACT

pCMV-NS35

+2 P S H I T A E A A G R R L A R G S P P S V A S S A  
4801 TCCCTCCCAT ATACAGCAG AGCGGCGCGG CGCAAGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA  
AGGAGGGTA TATTGTCGTC TCCGCCGCGC CGCTTCCAAC CGTCCCTA GTGGGGGAG ACACGGTGC AGGAGCGGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTGACCGC TAACCATGAC TCCCTGATG CTGAGCTCAT AGAGGCCAAC  
CGGTGGATAG GCGAGGTAGA GAGTTCGTT GAACGTGGG ATTGTA CTG AGGGACTAC GACTCGAGTA TCTCCGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
4961 CTCCTATGGA GCGAGGAGT GGGCGGCAAC ATCACCAGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTGA  
GAGGATACCT CCGTCTCTA CCGCGCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTTCAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
5041 TCCGCTTGTG GCGAGGAGG ACGAGCGGA GATCTCCGTA CCGCAGAAA TCCTGGGAA GTCTGGAGA TTCGCCAGG  
AGGGAACAC CGCTCCTCC TGCTCGCCT CTAGAGCAT GGGGCTTT AGGACGCTT CAGAGCCTCT AAGCGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
5121 CCCTGCCCCGTTGGGCGCGG CCGGACTATA ACCCCCGCT AGTGGAGACG TGGAAGAGC CCGACTACGA ACCACCTGTG  
GGGACGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCA TCACCTCTGC ACCTTTTCG GGCTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
5201 GTCCATGGCT GCGCGCTCC ACCTCCAAAG TCCCTCCTG TGCCTCGCC TCGGAAGAAG CGGACGGTGG TCCTCACTGA  
CAGGTACCGA CCGGCGAAG TGGAGTTTC AGGGAGGAC ACGGAGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
5281 ATCAACCCTA TCTACTGCCT TGGCGGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGCGACA  
TAGTTGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

pCMV-NS35

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
5361 ATACGACAAAC ATCCTCTGAG CCGGCCCTT CTGGCTGCC CCGGACTCC GACGCTGAGT CCTATTCTC CATGCCCC  
TATGCTGTG TAGGAGACTC GGGCGGGGAA GACCGACGG GGGGCTGAG CTGGACTCA GGATAAGGAG GTACGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
BamHI

5441 CTGGAGGGG AGCCTGGGA TCCGGATCTT AGCGACGGT CATGGTCAAC GGTCACTAGT GAGGCCAACG CGGAGGATGT  
GACCTCCCC TCGACCCCT AGGCTAGAA TCGCTGCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I  
5521 CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGGCGCACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA  
GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGCA CGCGGCGCT TCTTCTCTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K  
5601 ATGCACTAAG CAACCTGTTG CTACGTCACC ACAATTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGGCAGAAG  
TAGGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGGT CACGAACGGT TTCCGTCCTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K  
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA  
TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTGGTAA TGGTCTGCA TGAGTTCTC CAATTCTGC GCCGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y  
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACG CTGACGCCC CACACTCAGC CAAATCCAAG TTTGGTTATG  
TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTG GACTGGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

FIG. 3-Page 10

pCMV-NS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N  
5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT  
CCCGTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTIG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGTTC GTAAGCCAGC  
CATGTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCCAG CATTCGGTGC

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
6001 TCGTCTCATC GTGTTCCCG ATCTGGGCGT GCGCGTGC GAAAAGATGG CTTGTACGA CGTGGTTACA AAGCTCCCCT  
AGCAGAGTAG CACAAGGGC TAGACCCGCA CCGGCACACG CTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
EcoRI  
-----

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGAAGC GTGGAAGTCC  
ACGGGCACTA CCCTTCGAG ATCCCTAAGG TTATGAGTGG TCCTGTCGCC CAACTTAAGG AGCACGTTGC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
6161 AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA  
TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCGTGTAGG CATGCCTCCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
6241 GGCATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGG  
CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCCGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
6321 CCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCGCG GCGAGCGCGG TACTGACAAC TAGCTGTGGT  
CGGGAAGATG GTTAAGTTCC CCCCTCTGA CCGCGATAGC GTCCACGGCG CGCTCGCCCG ATGACTGTTG ATCGACACCA

FIG. 3-Page 11

PCMV-NS35

+2 N A T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
6401 AACACCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGACG GCTCCAGGAC TGCACCATGC TCGTGTGTGG  
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGGGTCC CGAGGTCCTG ACGTGGTAGG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGTCCAG GAGGACGGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA  
GCTGCTGAAT CAGCAATAGA CACTTTCGG CCCCCAGGTC CTCCTGGGCC GCTCGGACTC TCGGAAGTGC CTCGGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
6561 CCAGGTACTC CGCCCCCCT GGGACCCCC CACAACCAGA ATACGACTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
GGTCCATGAG GCGGGGGGA CCCCTGGGG GTGTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
6641 TCAGTCGCCC ACGACGGCG TGGAAAGAGG GTCTACTACC TCACCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG  
AGTCAGCGGG TGCTGCCGG ACCTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA  
CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
6801 TACTGATGAC CCATTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGGGAGAT CTACGGGGCC  
ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCCT GTTCGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCCG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
6881 TGCTACTCCA TAGAACCCT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA  
ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTAGTAAG TTTCTGAGT ACCGGAGTCC CGTAAAGTG AGGTGTCAAT

FIG. 3-Page 12

pCMV-NS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
6961 CTCTCCAGGT GAAATCAATA GGTGGGCGC ATGCCTCAGA AAACCTGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG  
GAGAGGTCCA CTTTAGTTAT CCCACGGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
7041 CCCGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCG GGTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA  
GGCCCTCGCA GCGCGGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCGGTGGC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
TCTTGTTTCG AGTTGAGTG AGTTATCGC CGGCGACCG TCGACCTGAA CAGGCCGACC AAGTCCGAC CGATGTCGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V  
7201 GGGAGACATT TATCACAGCG TGTCATGCG CCGGCCCGCG TGGATCTGGT TTTGCCCTACT CCTGCTTGT GCAGGGGTAG  
CCCTCTGTAA ATAGTCTCG ACAGATACG GGCCCGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R  
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGTAAAC ACTCCGGCCT AAAAAAATA AAAAAATCTAG AAAGCGCGCG  
CGTAGATGGA GGAGGGGTTG GCTACTTCCA ACCCCATTG TCGAGGCCGA TTTTITTTTTT TTTTITAGATC TTTCCGCGCG

BAMHI MluI  
-----

7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCCTTCT AGTTGCCAGC CATCTGTTGT  
GTTCTATAGT TCCTAGGTGA TCGCAATCT CGAGCGACTA CTCGGAGCTG ACACGGAAGA TCAACGGTGG GTAGACAACA

7441 TTGCCCCCTCC CCCGTGCCTT CCTTGACCTT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAATGAG GAAATTGCAT  
AACGGGAGG GGCACGGAA GGAAGTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGCAT TATTTTACTC CTTTAACGTA

pCMV-NS35

```

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTCT TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT
GGGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGCTTCC CCCTCCTAAC CCTTCTGTTA

7601 AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC CTTGGTTCGT TCGGCTGCGG CGAGCGGTAT
TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTCAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCCCTTCT TGTACACTCG TTTTCCGGTC

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGGG TTGCTGGCGT TTTTCCATAG GCTCGGCCCC CCTGACGAGC ATCACA AAAA
GTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGA CTGCTCGG TAGTGTTTTT

7841 TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATAAC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC
AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTTGAT ATTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCAGC

7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
CGAGAGGACA AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG

8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCCGT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCCGA
AGTGGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACCTG CTTGGGGGGC AAGTCGGGCT

8081 CCGCTGCGCC TTATCCGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
GGGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGTGGC CTAACACTAG CTACACTAGA
CATTGTCCCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

```



pCMV-NS35

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA  
TCCTGTCTATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT

---

8321 AACCAACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT  
TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTGCTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTAGGAA

---

8401 TGATCTTTTC TACGGGGTCT GAGGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTITGG TCATGAGATT ATCAAAAAAGG  
ACTAGAAAAG ATGCCCCAGA CTGGGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAAACC AGTACTCTAA TAGTTTTTCC

---

8481 ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTAA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG  
TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTTGA CCAGACTGTC

---

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGGGATC TGTCTATTTC GTTCATCCAT AGTTGCCGTA CTCCCCGTCG  
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

---

8641 TGAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGTGCA ATGATACCGC GAGACCCACG CTCACCGGCT  
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTACAGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

---

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA  
GGTCTAAATA GTCGTTATTT GGTCGGTCCG CCTTCCCGGC TCGGCTCTTC ACCAGGACGT TGAATAGGC GGAGGTAGGT

---

8801 GTCTATTAAAT TGTGCCCCGG AAGCTAGAGT AAGTAGTTCC CCAGTTAATA GTTTGGGCAA GGTGTGTGCC ATTGTACAG  
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGGCCTT GCAACAACGG TAACGATGTC

---

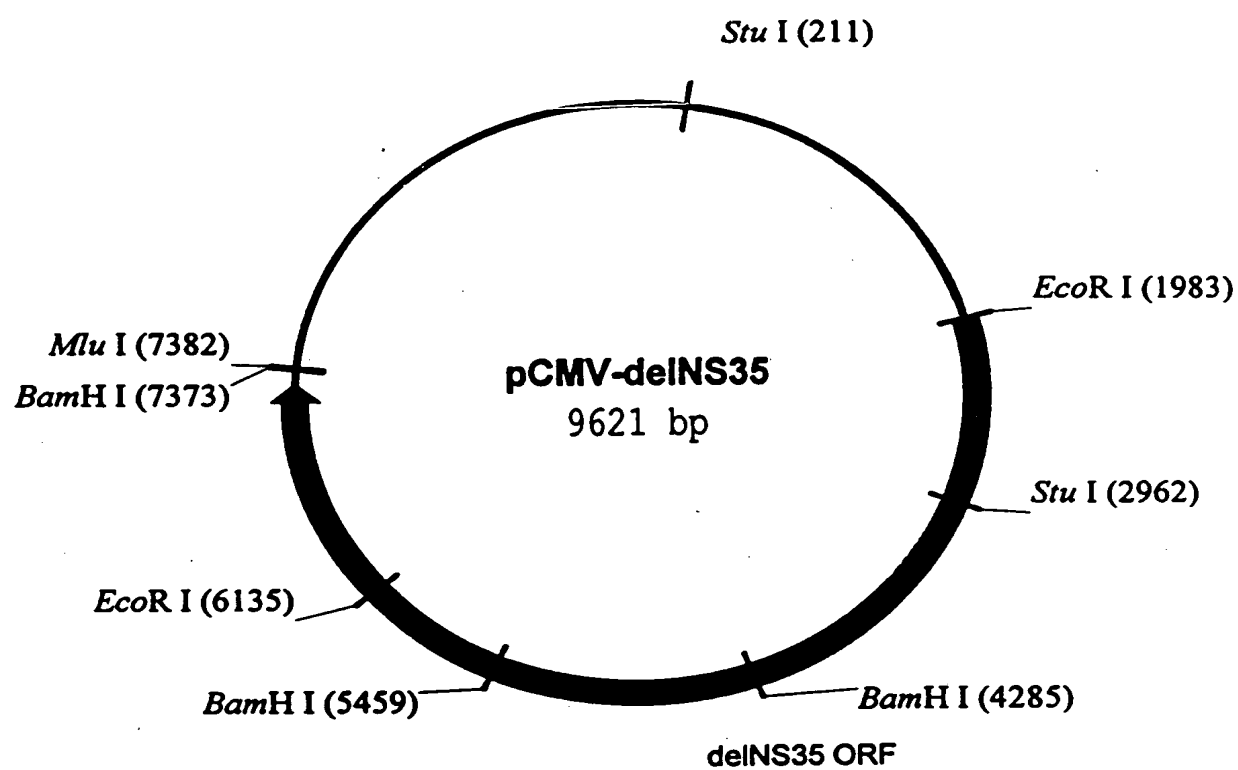
8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGCGGAGT TACATGATCC  
CGTAGCACCA CAGTCGGAGC AGCAAAACCAT ACCGAAGTAA GTCGAGGCCA AGGTTTGCTA GTTCCGCTCA ATGTACTAGG

---

FIG. 3-Page 15

pCMV-NS35

8961	CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTGGGTCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT GGGTACAACA CGTTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGCGTC ACAATAGTGA
9041	CATGGTTATG GCAGCACTGC ATAAATCTCT TACTGTCA TG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT
9121	CCAAGTCATT CTGAGAATAG TGTATGCGG GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT GGTTCAGTAA GACTCTTATC ACATACGCGG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA
9201	AGCAGAACTT TAAAAGTGCT CATCATTTGA AAACGTTCTT CGGGGCGGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG
9281	CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGTC GCAAAGACCC ACTCGTTTTT
9361	CAGGAAGCA AAATGCCGCA AAAAAGGGAA TAAGGGGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT GTCCCTTCGGT TTTACGGCGT TTTTTCCTT ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA
9441	TATTGAAGCA TTTATCAGG TTTATTGCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTTG TTTATCCCCA
9521	TCGGGCGACA TTTCCCGGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG
9601	GTATCAGGAG GCCCTTTTCGT C CATAGTGCTC CGGGAAAGCA G



**FIG. 4**

pCMV-delNS35

```

1  TCGCGCGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCTCT GTAAGCGGAT
   AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCGCTA

81  GCGGGGAGCA GACAAGCCCG TCAGGGGCGG TCAGGGGGTG TTGGCGGGTG TCGGGGGTGG CTTAACTATG CGGCATCAGA
   CGGCCCTCGT CTGTTGGGGC AGTCCCGCGC AGTCGCCAC AACCGGCCAC AGCCCCGACC GAATTGATAC GCGGTAGTCT

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTCGA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG
   CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGCCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG ATTGGGCGGA
   TTATCGAGTC TCGGGCTCGG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCGGCCTC TTACCCGCCT

321 ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACCTTGTATC TATATCATAA TATGTACATT TATATTGGCT
   TGACCCGCCC CTCCTTAAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTTAT
   GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

481 AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGCCCCATT
   TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTC ATTGACGTCA ATGGGTGGAG TATTACGGT
   CTGAGTTAT TACTGCATAC AAGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

```

StuI

-----

pCMV-deINS35

641 AAAC TGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC  
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

---

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGA CTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGCTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCAGGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGT

---

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAAACCCGC CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

---

961 CAAATGGCG GTAGCGGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

---

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCGG GGAACGCTGC ATTGGAACGC  
GGTAGGTGCG ACAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCGCGC CCTTGGCAGC TAACCTTGCG

---

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGCG ACGGTTCTCA CTGCATTTCAT GCGGATATC TGAGATATCC GTGTGGGGA ACCGAGAATA CGTACGATAT

---

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA  
GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

---

1281 TTGACCATT TTAGCCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACCTAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

pCMV-delINS35

1361	CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT	
	GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTCTACCC CAGGTAAATA	
1441	TATTACAAA TTCACATATA CAACAACGCC GTCCCGCGTG CCGGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT	
	ATAAATGTTT AAGTGTATAT GTTGTTCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCCCT AGAGGCTGTA	
1521	CTCGGGTAGG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GCGGAGCTT CCACATCGGA GCCCTGGTCC CATCCGTCCA	
	GAGCCCATGC ACAAGGCCGT TACCCGAGAA GAGGCCATCG CCGCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT	
1601	CGCGGTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC	
	CGCCGAGTAC CAGCGAGCGG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG	
1681	AGTGTGCGGC ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT	
	TCACACGGCG TGTTCGGCA CCGCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA	
1761	GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT	
	CCTTCTGAAT TCCGTGCGCG TCTTCTTCTA CTTCCGTCGA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA	
1841	TGCGGTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT	
	ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA	
		<div> <div>+2</div> <div> <div>M A A</div> <div>EcoRI</div> <div>----</div> </div> </div>
1921	GACAGACTAA CAGACTGTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT	
	CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA	

FIG. 5-Page 3

pCMV-delNS35

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K  
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTGCTGCA ACACGGGCT TTGGTGCTTA CATGTCCAAG  
TAGTCCGAGT CCCGATATTC CACGATCATG AGTTGGGAG ACAACGACGT TGTGACCCGA AACACCGAAT GTACAGGTTC

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G  
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCACGTACT CCACCTACGG  
CGAGTACCCT AGCTAGGATT GTAGTCCTGG CCCCACTCTT GTTAATGGTG ACCGTGCGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A  
2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGG GGGGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA  
GTTCAAGGAA CGGTGCGCG CCACGAGGCC CCGCGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T  
2241 CATCCATCTT GGGCATTGGC ACTGTCTTG ACCAAGCAGA GACTGCGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC  
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G  
2321 CCTCCGGGCT CCGTCACTGT CCCCATCCC AACATCGAGG AGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG  
GGAGGCCCCA GGCAGTGACA CCGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGCTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L  
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAA GAAGAAGTGC GACGAACTCG  
GTTCCGATAG GGGAGCTTC ATTAGTTCCC CCCCTCTGA GAGTAGAAGA CAGTAAGTTT CTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G  
2481 CCGCAAAGCT GTTCGCATTG GGCATCAATG CCGTGGCCTA CTACCGGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC  
GGCGTTTCCA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGGCGCA GAACTGCACA GGCAGTAGGG CTGGTCGCCC

FIG. 5-Page 4

pCMV-delNS35

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C  
2561 GATGTTGTG TCGTGGCAAC CGATGCCCTC ATACCGGCT ATACCGGCGA CTTGCACTCG GTGATAGACT GCAATACGTG  
CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S  
2641 TGTACCCAG ACAGTCGATT TCAGCCTTGA CCTACCTTC ACCATTGAGA CAATCAAGCT CCCCCAAGAT GCTGTCTCCC  
ACAGTGGGTC TGTACGCTAA AGTCGGAAT GGGATGGAAG TGGTAACCTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G  
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC  
CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCGG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V  
2801 ATGTTGCACT CGTCCGTCCT CTGTGAGTGC TATGAGCGAG GCTCTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT  
TACAACTGA GCAGGCAGGA GAACTCAGG ATACTCGTC CGACACGAAC CATACTCGAG TCGGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T  
StuI --

2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG  
ATCCGATGCT CCGATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCCCTCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q  
StuI ----

2961 GCCTCACTCA TATAGATGCC CACTTTCCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA  
CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGGTT

FIG. 5-Page 5



pCMV-delINS35

```

+2  A T V C A R A Q A P P S W D Q M W K C L I R L K P T
3041  GCCACCGTGT GCGTAGGC TCAAGCCCT CCCCACGT GGGACCAGT GTGGAAGTGT TTGATTGCGC TCAAGCCCAC
      CGGTGGCACA CGCGATCCCG AGTTCCGGGA GGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTGGGGTG

+2  L H G P T P L L Y R L G A V Q N E I T L T H P V T K
3121  CCTCCATGG CCAACACCC TGCTATACAG ACTGGCGCT GTTCAGAATG AAATCACCTT GAGGCACCCA GTCACCAAAT
      GGAGGTACCC GGTGTGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA

+2  Y I M T C M S A D L E V V T S T W V L V G G V L A A L
3201  ACATCATGAC ATGCATGTC GCGACCTGG AGTCTGTCAC GAGCACCTGG GTGCTCGTTG GCGCGTCTCT GCGTGTTCG
      TGTAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCAGGA CCGACGAAAC

+2  A A Y C L S T G C V V I V G R V V L S G K P A I I P D
3281  GCGCGGTATT GCCTGTCAAC AGGTGCGTG GTCATAGTGG GCAGGTCTGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA
      CGCGGCATAA CGGACAGTTG TCCGACGGAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGCCCCGT AGTATGGACT

+2  R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
3361  CAGGGAAGTC CTCTACCGAG AGTTCCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
      GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGTCTGTT CCTACTACG

+2  L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
3441  TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGTTATCGC CCTGTCTGTC
      AGCGGCTCGT CAAGTTCTGC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2  Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
3521  CAGACCAACT GGCAAAACCT CGAGACCTTC TGGCGGAAGC ATATGTGAA CTTCATCAGT GGGATACAAT ACTTGGCGGG
      GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCTATGTTA TGAACCGCCC

```

FIG. 5-Page 6

pCMV-delINS35

```

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
3601 CTTGTCAACG CTGCTGGTA ACCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACGAAAT GTGACGACA GTGTCGGGT GATTGGTGAT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
3681 GCCAAACCCT CCTCTTCAAC ATATTGGGG GGTGGGTGC TGCCAGCTC GCCCCCCCG GTGCCGCTAC TGCCTTTGTG
CGGTTTGGGA GGAGAAGTTG TATAACCC CACCCACCG ACGGTGCGAG CGCGGGGGG CACGGCGATG ACGAAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
3761 GCGCTGGCT TAGCTGGCG CGCCATCGG AGTGTGGAC TGGGAAGT CCTCATAGAC ATCCTTGCAG GGTATGGCG
CCGGGACCGA ATCGACCGG GCGGTAGCG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
3841 GGGCTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGTGCCT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
3921 CGCCATCCT CTCGCCGGA GCCCTCGTAG TCGGCGTGGT CTGTGACGA ATACTGCGC GGCAGCTTG CCGGGCGGAG
GGCGGTAGGA GAGCGGCCT CGGGAGCATC AGCGCACCA GACAGCTCT TATGACGCG CGTGCAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
4001 GGGCAGTGC AGTGGATGAA CCGCTGATA GCCTTCGCCT CCCGGGGAA CCATGTTTC CCCACGCACT ACGTCCCGGA
CCCCGTCAG TCACCTACTT GGCCGACTAT CGGAAGCGGA GGGCCCCCTT GTACAAAGG GGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
4081 GAGCGATGA GCTGCCCCG CTACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGTCTC GAGGCGACTG CACCAGTGA
CTCGCTACGT CGACGGGCG AGTGACGGA TGAGTCGTCG GAGTGACATT GGTGCGAGGA CTCGCGTGAC GTGGTCACCT

```

FIG. 5-Page 7

pCMV-delINS35

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D  
4161 TAAGCTCGA GTGTACCACT CCATGCTCG GTTCCTGGT AAGGACATC TGGACTGGA TATGCGAGGT GTTGAGCGAC  
ATTGAGCCT CACATGGTGA GGTACGAGG CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATAGGTTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G  
BamHI  
-----

4241 TTTAAGACCT GGCTAAAAGC TAAGTCAATG CCACAGCTGC CTGGGATCCC CTTTGTCTCC TGCAGCGCG GGTATAAGGG  
AAATCTGGA CCGATTTCG ATTGAGTAC GGTGTCGAG GACCCTAGGG GAAACACAGG ACGTCGCGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
4321 GGTCTGGCA GGGACGGCA TCATGCACAC TCGTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA  
CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
4401 TCAGGATCGT CGGTCCTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAAATG CCTACACCCAC GGGCCCCCTGT  
ACTCCTAGCA GCCAGGATCC TGGACGTCCT TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
4481 ACCCCCTTC CTGCGCGGAA CTACACGTTT GCGCTATGGA GGTGTCTGC AGAGGAATAC GTGCAGATAA GGCAGGTGG  
TGGGGGAAG GACGGGGCTT GATGTGCAAG CCGGATACCT CCCACAGAGG TCTCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCCGT GCCAGGTCCC ATCGCCCGAA TTTTTCACAG  
CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CCGTCCAGG TAGCGGGCTT AAAAAAGTGC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
4641 AATTGGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCTTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA  
TTAACCTGCC CCACGGGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

FIG. 5-Page 8

pCMV-delINS35

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
4721 CTCACGAAT ACCGGTAGG GTCCGAATTA CCTTGGAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA  
GAGTGTCTTA TGGGCCATCC CAGCGTAAAT GGAACGCTCG GCCTTGGCCT GCACCGGCAC AACTGCAGT ACGAGTGA

+2 P S H I T A E A A G R R L A R G S P P S V A S S A  
4801 TCCCTCCCAT ATAACAGCAG AGGGGCGG GCGAAGTTG GCGAGGGAT CACCCCTC TGTCGCCAGC TCCTCGGCTA  
AGGAGGGTA TATTGTCGTC TCCGCCGCC CGCTTCCAAC CGCTCCCTA GTGGGGGAG ACACCGGTG AGGAGCCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC  
CGGTGATAG GCGAGGTAGA GAGTCCGTT GAACGTGGCG ATTGTAAGT AGGGGACTAC GACTCGAGTA TCTCCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
4961 CTCCTATGGA GGCAGGAGAT GGGGGGCAAC ATCACCAGG TTGACTCAGA AAACAAAGT GTGATTCGTG ACTCCTTCCA  
GAGGATACCT CCGTCTCTA CCGCGCGTTG TAGTGTCCC AACTCAGTCT TTTGTTTAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
5041 TCCGCTTGTG GCGAGGAGG ACGAGCGGA GATCTCCGTA CCGCAGAAA TCCTGGGAA GTCTGGAGA TTCGCCCCAGG  
AGGCGAACAC CGCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTT AGGACGCCCT CAGAGCCTCT AAGCGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
5121 CCCTGCCGT TTGGGCGG CCGGACTATA ACCCCCGCT AGTGAGAGC TGGAATAAG CCGACTACGA ACCACCTGTG  
GGGACGGCA AACCCGCGC GCCTGATAT TGGGGGCGA TCACCTCTGC ACCTTTTCG GGTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
5201 GTCCATGGCT GCCGCTTCC ACCTCAAAG TCCCTCCTG TCCCTCCGC TCGAAGAAG CGGACGGTGG TCCTCACTGA  
CAGTACCGA CCGGCGAAG TGGAGTTTC AGGGAGGAC ACGGAGCGG AGCTTCTTC GCCTGCCACC AGGAGTGA

pCMV-deINS35

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
5281 ATCAACCCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGCGGGACA  
TAGTTGGGAT AGATGACGGA ACGGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTT CTGGCTGCC CCCGACTCC GACGCTGAGT CCTATTCTC CATGCCCCC  
TATGCTGTTG TAGGAGACTC GGGGGGGAA GACCGACGG GGGGCTGAG CTGCGACTCA GGATAAGGAG GTACGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
BamHI  
-----

5441 CTGGAGGGG AGCCTGCGCA TCCGATCTT AGCGACGGT CATGTCAC GGTCACTAGT GAGGCCAACG CGGAGGATGT  
GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCA GTACCAGTTG CCAGTCATCA CTCGGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I  
5521 CGTGTGCTGC TCAATGTCTT ACTCTTGAC AGGGCACTC GTCACCCCGT GCGCGCGGA AGAACAGAAA CTGCCCCATCA  
GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGCA GCGGGCCCT TCTTGTCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K  
5601 ATGCACTAAG CAACTCGTTG CTACGTACC ACAATTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGCAGAAG  
TAGGTGATTC GTGAGCAAC GATGACGTGG TGTTAAACCA CATAAGTGG TGGAGTGGT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K  
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA  
TTTCAGTGTA AACTGTCTGA GGTCAAGAC CTGTCCGTAA TGGTCTGCA TGAGTTCTC CAATTCTC GCGGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y  
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACG CTGACGCCCC CACACTCAGC CAAATCCAAG TTGTTATG  
TCACTTCCGA TTGAACGATA GGCATCTCT TCGAACGTG GACTCGGGG GTGTGAGTGG GTTAGGTTT AAACCAATAC

FIG. 5-Page 10

pCMV-deINS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N  
5841 GGGCAAAAGA CGTCCGTTGC CATGCAGAA AGCCGTAAC CCACATCAAC TCCGTGTGA AAGACCTTCT GGAAGACAAT  
CCCGTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC  
CATTTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCAG CATTCGGTCC

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
6001 TCGTCTCATC GTGTTCCCG ATCTGGCGT GCGCGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT  
AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGGCACAG CTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
EcoRI

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC  
ACCGGCACTA CCCTCGAGG ATGCCTAAGG TTATGATGG TCCTGTCGCC CAACTTAAGG AGCACGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
6161 AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGGCACATCC GTACGGAGGA  
TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGGAAGA AACTGAGGTG TCAGTGAATC TCGCTGTAGG CATGCCTCCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG  
CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGCGG CACCGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT CGGGCTATCG CAGGTGCCG GCGAGGGCG TACTGACAAC TAGCTGTGGT  
CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCGGATAGC GTCCACGGCG CGCTCGCCG ATGACTGTTG ATCGACACCA

FIG. 5-Page 11

pCMV-delINS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
6401 AACACCCTCA CTTGCTACAT CAAGGCGCGG GCAGCCTGTC GAGCGGCAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG  
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTTG ACGTGGTACG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGTCCAG GAGGACGCGG CGAGCCTCAG AGCCTTCACG GAGGCTATGA  
GCTGCTGAAT CAGCAATAGA CACTTTCCGG CCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
6561 CCAGGTACTC CGCCCCCCT GGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
GGTCCATGAG GCGGGGGGGA CCCCTGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
6641 TCAGTCGCCC ACGACGGCGC TGGAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGGA GAGTGGGTG  
AGTCAGCGGG TGCTGCGCGG ACCTTCTCC CAGATGATGG AGTGGCACT GGGATGTTGG GGGGAGCGCT CTGACGCGAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTCGCC CCACACTGTGG GCGAGGATGA  
CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
6801 TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGGGAGAT CTACGGGGCC  
ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCCT GGTGCAACTT GTCCGGGAGC TAACGCTCTA GATCCCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
6881 TGCTACTCCA TAGAACCACCT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA  
ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTAGTAAG TTTCTGAGT ACCGGAGTGC CGTAAAGTG AGGTGTCAAT

pCMV-delINS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
6961 CTCTCCAGGT GAAATCAATA GGTGGCCGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACCCGG  
GAGAGGTCCA CTTTAGTTAT CCCACCGCG TACGGAGTCT TTTGAACCCC ATGGCGGGA CGCTCGAACC TCTGTGCCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
7041 CCCGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGCGAG GGTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA  
GGGCTCGCA GCGCGGATCC GAAGACCGGT CTCCTCCGC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
TCTTGTTTCG AGTTTGAGTG AGTTATCGC CGCGGACCG TCGACCTGAA CAGGCCGACC AAGTCCCGAC CGATGTCGCC

+2 G D I Y H S V S H A R P R W I W F C L L L A A G V  
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCGC TGGATCTGCT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG  
CCCTCTGTAA ATAGTCTCG ACAGAGTACG GGCCGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R  
7281 GCATCTACCT CCTCCCCAAC CGATGAAGT TGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAATCTAG AAAGGCGCGC  
CGTAGATGGA GGAGGGGTG GCTACTTCCA ACCCCATTG TGAGGCCGGA TTTTITTTTT TTTTITAGATC TTTCCGCGCG

BamHI MluI  
-----

7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT  
GTTCTATAGT TCCTAGGTGA TCGGCAATCT CGAGCGACTA CTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA

7441 TTGCCCCCTCC CCCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACCT TCCTTTCCTA ATAAATGAG GAAATTGCAT  
AACGGGGAGG GGGCACGGAA GGAAGTGGGA CCTTCCACCG TGAGGGTGAC AGGAAAGCAT TATTTTACTC CTTTAAACGTA



pCMV-deINS35

7521	CGCATTGTCT GAGTAGGTGT CATTCTATTTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGGTTCC CCCTCCTAAC CCTTCTGTTA
7601	AGCAGGCATG CTGGGGAGCT CTTCGGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTTCGT TCGGCTGCGG CGAGCGGTAT TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGACTGA CTGAGCGAGC CGAGCCAGCA AGCCGACGCC GCTCGCCATA
7681	CAGTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG GTCGAGTGAG TTTCCGGCCAT TATGCCAATA GGTGTCTTAG TCCCCATTG CGTCCTTTCT TGTACACTCG TTTTCCGGTC
7761	CAAAAGGCCA GGAACCGTAA AAAGGCCGGG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAATAA GTTTTCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGAAGTCTCG TAGTGTTTT
7841	TCGACGCTCA AGTCAGAGGT GCGGAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTTGAT ATTCTATGG TCCGCAAAGG GGGACCTTCG AGGAGCACG
7921	GCTCTCTGT TCCGACCCCTG CCGCTTACCG GATACCTGC CGCTTTTCTC CCTTCGGGA GCGTGGCGCT TTCTCAATGC CGAGAGGACA AGGTGGGAC GCGGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG
8001	TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCCGT CCAAGCTGGG CTGTGTGCAC GAACCCCGCG TTCAGCCCCA AGTCCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT
8081	CCGCTGGCC TTATCCGGA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GGGACGGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC
8161	GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACTAGG CTACACTAGA CATTGTGCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATCTGATCT

FIG. 5-Page 14

pCMV-deINS35

8241	AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA TCCTGTGCATA AACCATAGAC GCGAGACGAC TTGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT
8321	AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TTGGTGGCGA CCATCGCCAC CAAAAAACA AAGTTTCGTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTTAGGAA
8401	TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCTAAAAACC AGTACTCTAA TAGTTTTTCC
8481	ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG TAGAAGTGGA TCTAGGAAAA TTTAATTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTTGA CCAGACTGTC
8561	TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGGGATC TGCTATTTC GTTCATCCAT AGTTGCCGTA CTCCTCCGTCG AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGGTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC
8641	TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGTGCA ATGATACCGC GAGACCCACG CTCACCGGCT ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG CTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA
8721	CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCG AGCGCAGAAG TGGTCTTCCA ACTTTATCCG CCTCCATCCA GGTCTAAATA GTCGTTATT GGTGGGTGG CCTTCCCGGC TCGGTCTTC ACCAGGACGT TGAATAGGC GGAGGTAGGT
8801	GTCTATTAAT TGTGCGCGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGGGCAA CGTTGTTGCC ATTGCTACAG CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGGCTT GCAACAACGG TAACGATGTC
8881	GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC CGTAGCACCA CAGTGGGAGC AGCAAAACCAT ACCGAAGTAA GTCGAGGCCA AGGTTGCTA GTTCCGCTCA ATGTACTAGG

pCMV-deINS35

```

8961  CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTGGTCTCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTATCACT
      GGGTACAACA CGTTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATT CACCGGCGTC ACAATAGTGA

9041  CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCA TG CCAATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA
      GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GTAGGGCATT CTACGAAAAG ACAC TGACCA CTCATGAGTT

9121  CCAAGTCATT CTGAGAATAG TGTATCGGC GACCGAGTTG CTCTTGCCCG CGGTCAATAC GGGATAATAC CGGCCACAT
      GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGC CGCAGTTATG CCCTATTATG CGCGGTGTA

9201  AGCAGAACTT TAAAGTGCT CATCATTTGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC
      TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281  CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
      GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGTC GCAAAGACCC ACTCGTTTTT

9361  CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
      GTCCTTCCGT TTTACGGCGT TTTTTCCTT ATTCCCGCTG TGCTTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441  TATTGAAGCA TTTATCAGG TTTATGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAAATAAC AAATAGGGGT
      ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTATG TTTATCCCCA

9521  TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCAATGAC ATTAACCTAT AAAAATAGGC
      AGGCGGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

9601  GTATCACGAG GCCCTTTCGT C
      CATAGTGCTC CGGAAAAGCA G

```

FIG. 5-Page 16

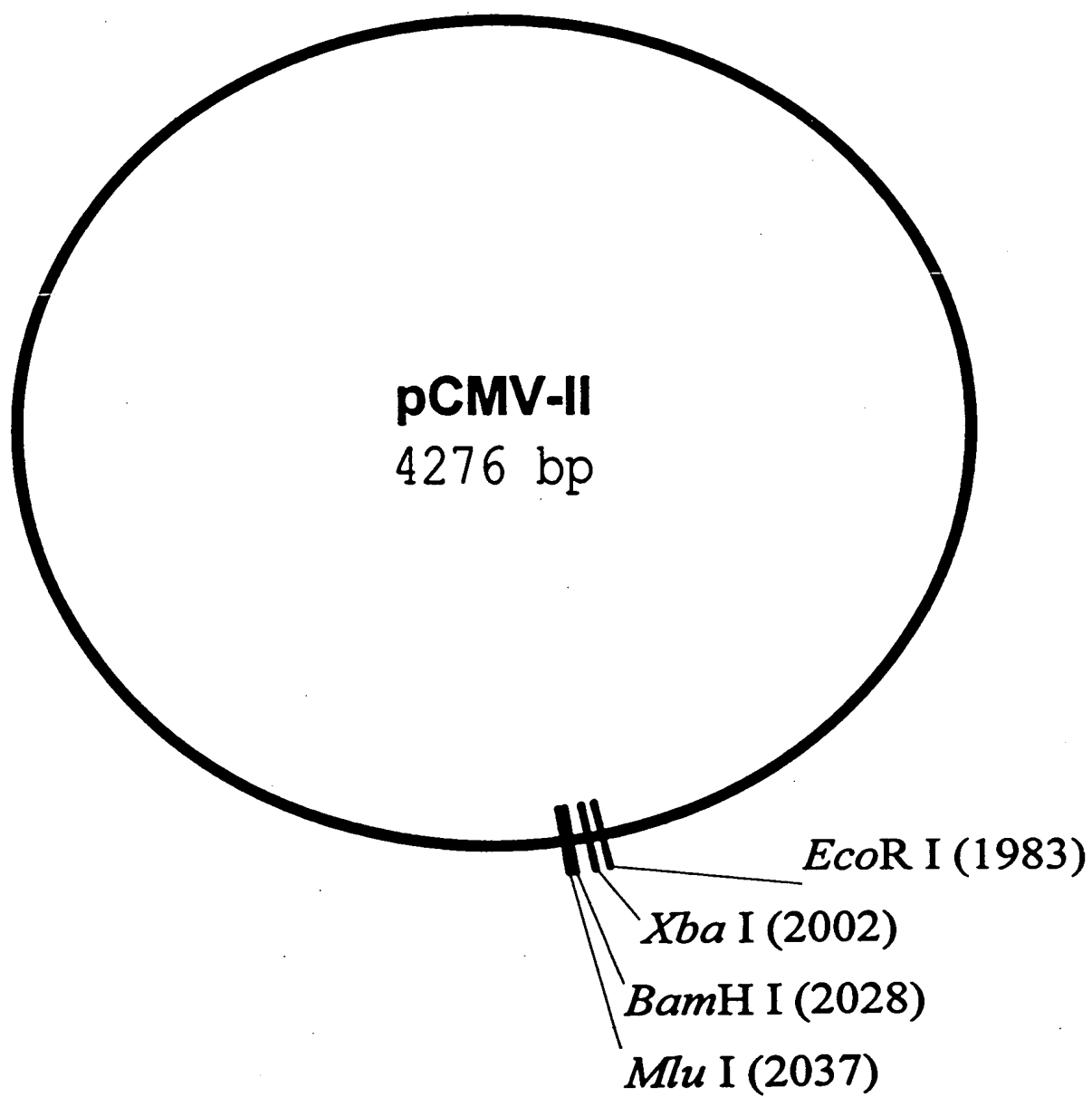


FIG. 6

pCMV-II

1	TCGCGCGCTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CCTCGAGGGC	CTCTGCCAGT	GTGGAACAGA	CATTCCGCTA
81	GC	CGGGGAGCA	GACAAGCCCG	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACTATG
	CGGCCCTCGT	CTGTTCCGGC	AGTCCCGCGC	AGTCCGCCAC	AACCGCCAC	AGCCCCGACC	GAATTGATAC	GCCGTAGTCT
161	GC	GAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTGTGA	AAAGCCTAGG	CCTCCAATAA	AGCCTCCTCA
	CGTCTAACAT	GACTCTCAG	TGGTATACTT	CGAAAAACGT	TTTCGGATCC	GGAGTTTTT	TCGGAGGAGT	GATGAAGACC
241	AATAGCTCAG	AGCCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAT	TAGTCAGCCA	TGGGGCGGAG	AATGGCGGGA
	TTATCGAGTC	TCCGGCTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT	ACCCCGCCTC	TTACCCGCCCT
321	ACTGGCGGG	GAGGGAATTA	TTGGCTATTG	GCCATTGCGT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	TGACCCGCCC	CTCCCTTAAT	AACCGATAAC	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT
	GTACAGGTTA	TACTGGCGGT	ACAACTGTAA	CTAATAACTG	ATCAATAATT	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA
481	AGCCCATATA	TGGAGTTCCG	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	CCCGCCCATTT
	TCGGGTATAT	ACCTCAAGGC	GCAATGTATT	GAATGCCAAT	TACCGGGCGG	ACCGACTGGC	GGGTTGCTGG	GGCGGGGTAA
561	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AAGGCCAATA	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT
	CTGCAGTTAT	TACTGCATAC	AAGGTATCA	TTGCGGTTAT	CCCTGAAAGG	TAACTGCAGT	TACCCACCTC	ATAAATGCCA
641	AAACTGCCCC	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCGGCCC	CCTATTGACG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTCAT	GTAGTTTACA	TAGTATACGG	TTCAGGGCGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG

FIG. 7-Page 1

pCMV-II

721	GCCTGGCAAT ATGCCCAGTA CATGACCTTA CGGGACCTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC CGGACCGTAA TACGGGTCAAT GTACTGGAAT GGCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG
801	CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGGCGTGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA GTACCACTAC GCCAAAACCG TCATGTGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGTTCA GAGGTGGGGT
881	TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTTCCA AAATGTGTA ATAACCCCGC CCCGTTGACG AACTGCAGT ACCCTCAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC
961	CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC
1041	CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGGCGCGG GGAACGGTGC ATTGGAACGC GGTAGGTGG ACAAAACTGG AGGTATCTTC TGTGCCCTG GCTAGGTGCG AGCGCGCGGC CCTTGCCACG TAACCTTGCG
1121	GGATTCCCG TGCCAAGAGT GACGTAAGTA CCGCTATAG ACTCTATAG CACACCCCTT TGGCTCTTAT GCATGCTATA CCTAAGGGC ACGGTTCTCA CTGCATTCTAT GCGGATATC TGAGATATCC GTGTGSGGAA ACCGAGAATA CGTACGATAT
1201	CTGTTTTGG CTGGGGCCCT ATACACCCCG GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT
1281	TTGACCATT TTAGCCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACAT AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA
1361	CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA

pCMV-II

1441	TATTTACAAA	TTCACATATA	CAACAACGCC	GTCCCCCGTG	CCCGCAGTTT	TTATTAAACA	TAGCGTGGGA	TCTCCGACAT
	ATAAATGTTT	AAGTGATAT	GTTGTTGCGG	CAGGGGGCAC	GGCGGTCAAA	AATAATTGT	ATCGCACCT	AGAGGCTGTA
1521	CTCGGGTACG	TGTTCCGGAC	ATGGGCTCTT	CTCCGGTAGC	GGCGGAGCTT	CCACATCCGA	GCCCTGGTCC	CATCCGTCCA
	GAGCCCATGC	ACAAGGCCTG	TACCCGAGAA	GAGCCCATCG	CCGCCTCGAA	GGTGTAGGCT	CGGGACCAGG	GTAGGCAGGT
1601	GCGGCTCATG	GTGGCTCGGC	AGCTCCTTGC	TCCTAACAGT	GGAGGCCAGA	CTTAGGCACA	GCACAATGCC	CACCACCACC
	CGCCGAGTAC	CAGCGAGCCG	TCGAGGAACG	AGGATTGTCA	CCTCCGGTCT	GAATCCGTCT	CGTGTACGG	GTGGTGGTGG
1681	AGTGTGCCGC	ACAAGGCCGT	GGCGGTAGGG	TATGTGTCTG	AAAATGAGCT	CGGAGATTGG	GCTCGCACCT	GGACGCAGAT
	TCACACGGCG	TGTTCCGGCA	CGGCCATCCC	ATACACAGAC	TTTTACTCGA	GCCTCTAACC	CGAGCGTGA	CCTGCCGTCTA
1761	GGAGACTTA	AGGCAGCGGC	AGAAGAAGAT	GCAGGCAGCT	GAGTTGTTGT	ATTCTGATAA	GAGTCAAGG	TAACTCCCGT
	CGTTCTGAAT	TCCGTCCCGC	TCITTCTTCTA	CGTCCGTCTGA	CTCAACAACA	TAAGACTATT	CTCAGTCTCC	ATTGAGGGCA
1841	TGCGGTGCTG	TTAACGGTGG	AGGGCAGTGT	AGTCTGAGCA	GTACTCGTTG	CTGCCGCCCG	CGCCACCAGA	CATAATAGCT
	ACGCCACGAC	AATTGCCACC	TCCCGTCACA	TCAGACTCGT	CATGAGCAAC	GACGGCGCGC	GCGGTGGTCT	GTATTATCGA
1921	GACAGACTAA	CAGACTGTTT	CTTTCCATGG	GTCTTTTCTG	CAGTCACCGT	CGTCGACCTA	AGAAATCAGA	CTCGAGCAAG
	CTGTCTGATT	GTCTGACAAG	GAAAGGTACC	CAGAAAAGAC	GTCAGTGGCA	GCAGCTGGAT	TCTTAAGTCT	GAGCTCGTTC
2001	TCTAGAAAGG	CGCGCCAAGA	TATCAAGGAT	CCACTACGGG	TTAGAGCTCG	CTGATCAGCC	TCGACTGTGC	CTTCTAGTTG
	AGATCTTTCC	GCGGGTTTCT	ATAGTTCTTA	GGTGATGGC	AATCTCGAGC	GACTAGTCGG	AGCTGACACG	GAAGATCAAC

EcoRI  
-----

XbaI                      BamHI                      MluI  
-----

pCMV-II

2081 CCAGGCATCT GTTGTCTTGGC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCTTT TCCTAATAAA  
GGTCGGTAGA CAACAAACGG GGAGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA AGCATTATTT

---

2161 ATCAGGAAAT TGCATCGCAT TGCTCGAGTA GGTCATTC TATTCTGGG GGTGGGTGG GGCAGGACAG CAAGGGGAG  
TACTCCCTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG ATAAGACCCC CCACCCACC CCGTCTGTG GTTCCCCCTC

---

2241 GATTGGGAAG ACAATAGCAG GCATGCTGG GAGCTCTTCC GCTTCCTCG TCACTGACTC GCTGCGCTCG GTCTTTCGGC  
CTAACCCCTC TGTATCGTC CGTACGACCC CTCGAGAAGG CGAAGGAGCG AGTGACTGAG CGACGCGAGC CAGCAAGCCG

---

2321 TCGGGCGAGC GGTATCAGCT CACTCAAAG CGGTAATACG GTTATCCACA GAATCAGGG ATAACGCAGG AAAGAACATG  
ACGGCGCTCG CCATAGTCGA GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTCTTTGTAC

---

2401 TCAGCAAAAG GCCAGCAAA GCCAGGAAC CGTAAAAAG CCGCGTTGCT GCGGTTTTTC CATAGGCTCC GCCCCCCTGA  
ACTCGTTTT CCGTCGTTTT CCGGTCCCTG GCATTTTTTC GCGCAACGA CCGCAAAAG GTATCCGAG CCGGGGACT

---

2481 CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGTGGCGA AACCCGACAG GACTATAAG ATACCAGGCG TTTCGCCCTG  
GCTCGTAGTG TTTTATAGCTG CGAGTTCAGT CTCCACCGCT TTGGGCTGTC CTGATAATTC TATGTTCCG AAAGGGGGAC

---

2561 GAAGCTCCCT CGTGCGCTCT CCTGTTCGA CCCTGCCGCT TACCGGATAC CTGTCCGCCCT TTCTCCCTTC GGGAAGCGTG  
CTTCGAGGGA GCACGCGAGA GAACAAGGCT GGCACGGCGA ATGGCCTATG GACAGGCGGA AAGAGGGAAG CCTTTCGCAC

---

2641 GCGTTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTCGG TGTAGGTCTG TCGTCCAAG CTGGGCTGTG TGCACGAACC  
CGCGAAAGAG TTACGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGGTTT GACCCGACAC ACGTCTTGG

---

2721 CCCCCTTCAG CCGACCGCT GCGCCTTATC CCGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC  
GGGGCAAGTC GGGCTGGGA CCGGGAATAG GCCATTGATA GCAGAACTCA GGTGGGCCA TTCTGTCTG AATAGCGGTG

---



pCMV-II

2801	TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC ACCGTCGTGG GTGACCATTG TCCTAATCGT CTGCTCCAT ACATCCGCCA CGATGCTCA AGAATTTCAC CACCGGATTG
2881	TACGGCTACA CTAGAAAGGAC AGTATTTGGT ATCTGGGCTC TGCTGAAGCC AGTTACCTTC GGA AAAAGAG TTGGTAGCTC ATGCCGATGT GATCTTCCTG TCATAAACCA TAGACGGGAG ACGACTTCGG TCAATGGAAG CCTTTTCTC AACCATCGAG
2961	TTGATCCGGC AAACAAACCA CCGCTGGTAG CCGTGGTTT TTTGTTTGCA AGCAGCAGAT TACGGCGAGA AAAAAGGAT AACTAGGCGG TTTGTTTGGT GCGGACCATC GCCACCAAAA AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTCCTA
3041	CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG GACTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTCAGTG CAATTCCTTA AAACCATAC
3121	AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAAT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGACTA TCTAATAGTT TTTCCCTAGAA GTGGATCTAG GAAAATTTAA TTTTACTTC AAAATTAGT TAGATTTCAT ATATACTCAT
3201	AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTTCGTTCA TCCATAGTTG TTGAACCCAGA CTGTCAATGG TTACGAATTA GTCACCTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC
3281	CCTGACTCCC CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT GGGCCCCAGTG CTGCAATGAT ACCGGGAGAC GGACTGAGGG GCAGCACATC TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTAC GACGTTACTA TGGCGCTCTG
3361	CCACGGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT GGTGGGAGTG GCGGAGTCT AAATAGTCTG TATTGGTGG GTCCGCTTC CCGGCTCGG TCTTCACCAG GACGTTGAAA
3441	ATCCGGCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG TAGCGGAGG TAGTCAGAT AATTAAACAC GGCCCTTCGA TCTCATTTCAT CAAGCGGTCA ATTATCAAAAC GCGTTGCAAC

FIG. 7 -Page 5

pCMV-II

3521 TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTACGCT CCGGTTCCCA ACGATCAAGG  
AACGGTAACG ATGTCCGTAG CACCACAGTG CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGT TGCTAGTTCC

---

3601 CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTCAAGG GTAAAGTTGGC  
GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTCGGCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCACACCG

---

3681 CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA  
GGTCACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAAGATGAC AGTACGGTAG GCATTCTACG AAAAGACACT

---

3761 CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT  
GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCGCTGGC TCAACGAGAA CGGCGCGCAG TTATGCCCTA

---

3841 AATACCGCGC CACATAGCAG AACTTTAAAA GTGTCATCA TTGGAAGACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT  
TTATGGCGCG GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC GCTTTTGAGA GTTCCTAGAA

---

3921 ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT  
TGGGACAAAC TCTAGGTCAA GCTACATTGG GTGAGCACGT GGGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

---

4001 CTGGGTGAGC AAAAAACAGG AGGCAAAATG CCGCAAAAAA GGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC  
GACCCACTCG TTTTGTGCTT TCCGTTTAC GCGGTTTTT CCCTTATTC CGCTGTGCCT TTACAACCTA TGAGTATGAG

---

4081 TTCCTTTTTC AATATTATTG AAGCATTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTAGAAAAA  
AAGGAAAAAG TTATAATAAC TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTTACAT AAATCTTTTT

---

4161 TAAACAAATA GGGGTTCCGC GCACATTTC CCGAAAAGTG CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA  
ATTGTGTTAT CCCCAGGCG CGGTAAAGG GGCTTTTCAC GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT

---

4241 CCTATAAAA TAGGCGTATC ACGAGGCCCT TTGCTC  
GGATATTTTT ATCCGCATAG TGCTCGGGA AAGCAG

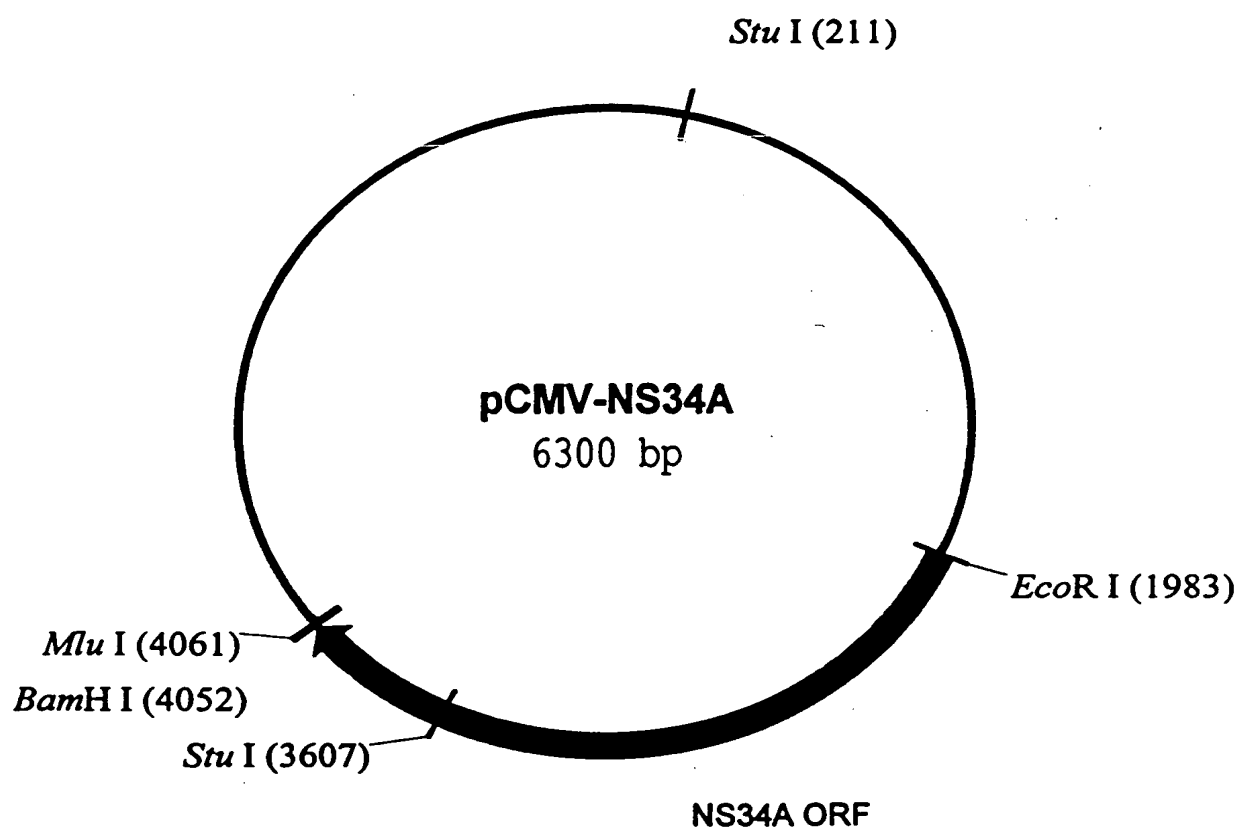


FIG. 8

pCMV-NS34A

1	TCGCGCGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCC
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC
51	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT	GCCGGGAGCA	GACAAGCCCC
	CTCTGCCAGT	GTCGAACAGA	CATTCGCCTA	CGGCCCTCGT	CTGTTCTGGG
101	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATATG
	AGTCCCAGCG	AGTCGCCAC	AACCGCCAC	AGCCCCGACC	GAATTGATAC
151	CGGCATCAGA	GCAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTGTGA
	GCCGTAGTGT	CGTCTAACAT	GACTCTCAG	TGGTATACTT	CGAAAAACGT
StuI ~~~~~					
201	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG	AATAGCTCAG
	TTTCGGATCC	GGAGGTTTTT	TCGGAGGAGT	GATGAAGACC	TTATCGAGTC
251	AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAAT	TAGTCAGCCA
	TCCGGCTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT
301	TGGGGCGGAG	AATGGGCGGA	ACTGGGCGGG	GAGGGAATTA	TGGGCTATTG
	ACCCCGCCTC	TTACCCGCCT	TGACCCGCCC	CTCCCTTAAT	AACCGATAAC
351	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA
	GTACAGGTTA	TACTGGCGGT	ACAACGTGTA	CTAATAACTG	ATCAATAATT
451	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT	AGCCCATATA	TGGAGTTCGG
	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA	TCGGGTATAT	ACCTCAAGGC
501	CGTTACATAA	CTTACGGTAA	ATGGCCCCGC	TGGCTGACCG	CCCAACGACC
	GCAATGTATT	GAATGCCATT	TACCGGGCGG	ACCGACTGGC	GGGTTGCTGG
551	CCCGCCCAT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA
	GGGCGGGTAA	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT
601	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT	AAACTGCCCA
	CCCTGAAAGG	TAACTGCAGT	TACCCACCTC	ATAAATGCCA	TTTGACGGGT
651	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCCC	CCTATTGACG
	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC
701	TCAATGACGG	TAAATGGCCC	GCCTGGCATT	ATGCCAGTA	CATGACCTTA
	AGTTACTGCC	ATTTACCGGG	CGGACCGTAA	TACGGGTCAT	GTACTGGAAT
751	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC
851	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG
	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT	AACTGCAGTT	ACCCTCAAAC

pCMV-NS34A

901	TTTTGGCACC AAAACCGTGG	AAAATCAACG TTTtagTTGC	GGACTTTCCA CCTGAAAGGT	AAATGTCGTA TTTACAGCAT	ATAACCCCGC TATTGGGGCG
951	CCCGTTGACG GGGCAACTGC	CAAATGGGCG GTTTACCCGC	GtagGCGTGT CATCCGCACA	ACGGTGGGAG TGCCACCCTC	GTCTATATAA CAGATATATT
1001	GCAGAGCTCG CGTCTCGAGC	TTTAGTGAAC AAATCACTTG	CGTCAGATCG GCAGTCTAGC	CCTGGAGACG GGACCTCTGC	CCATCCACGC GGTAGGTGCG
1051	TGTTTTGACC ACAAAACTGG	TCCATAGAAG AGGTATCTTC	ACACCGGGAC TGTGGCCCTG	CGATCCAGCC GCTAGGTCGG	TCCGCGGCCG AGGCGCCGGC
1101	GGAACGGTGC CCTTGCCACG	ATTGGAACGC TAACCTTGCG	GGATTCCCCG CCTAAGGGGC	TGCCAAGAGT ACGGTTCTCA	GACGTAAGTA CTGCATTTCAT
1151	CCGCCTATAG GGCGGATATC	ACTCTATAGG TGAGATATCC	CACACCCCTT GTGTGGGGAA	TGGCTCTTAT ACCGAGAATA	GCATGCTATA CGTACGATAT
1201	CTGTTTTTGG GACAAAACC	CTTGGGGCCT GAACCCCGGA	ATACACCCCC TATGTGGGGG	GCTCCTTATG CGAGGAATAC	CTATAGGTGA GATATCCACT
1251	TGGTATAGCT ACCATATCGA	TAGCCTATAG ATCGGATATC	GTGTGGGTTA CACACCCAAT	TTGACCATTA AACTGGTAAT	TTGACCACTC AACTGGTGAG
1301	CCCTATTGGT GGGATAACCA	GACGATACTT CTGCTATGAA	TCCATTACTA AGGTAATGAT	ATCCATAACA TAGGTATTGT	TGGCTCTTTG ACCGAGAAAC
1351	CCACAATAT GGTGTGATA	CTCTATTGGC GAGATAACCG	TATATGCCAA ATATACGGTT	TACTCTGTCC ATGAGACAGG	TTCAGAGACT AAGTCTCTGA
1401	GACACGGACT CTGTGCCTGA	CTGTATTTTT GACATAAAAA	ACAGGATGGG TGTCTTACCC	GTCCATTTAT CAGGTAAATA	TATTTACAAA ATAAATGTTT
1451	TTCACATATA AAGTGTATAT	CAACAACGCC GTTGTTGCGG	GTCCCCCGTG CAGGGGGCAC	CCCGCAGTTT GGGCGTCAAA	TTATTAAACA AATAATTGT
1501	TAGCGTGGGA ATCGCACCTT	TCTCCGACAT AGAGGCTGTA	CTCGGGTACG GAGCCCATGC	TGTTCCGGAC ACAAGGCCTG	ATGGGCTCTT TACCCGAGAA
1551	CTCCGGTAGC GAGGCCATCG	GGCGGAGCTT CCGCCTCGAA	CCACATCCGA GGTGTAGGCT	GCCCTGGTCC CGGGACCAGG	CATCCGTCCA GTAGGCAGGT
1601	GCGGCTCATG CGCCGAGTAC	GTCGCTCGGC CAGCGAGCCG	AGCTCCTTGC TCGAGGAACG	TCCTAACAGT AGGATTGTCA	GGAGGCCAGA CCTCCGGTCT
1651	CTTAGGCACA GAATCCGTGT	GCACAATGCC CGTGTTACGG	CACCACCACC GTGGTGGTGG	AGTGTGCCGC TCACACGGCG	ACAAGGCCGT TGTTCCGGCA
1701	GGCGGTAGGG CCGCCATCCC	TATGTGTCTG ATACACAGAC	AAAATGAGCT TTTTACTCGA	CGGAGATTGG GCCTCTAACC	GCTCGCACCT CGAGCGTGGA
1751	GGACGCAGAT CCTGCGTCTA	GGAAGACTTA CCTTCTGAAT	AGGCAGCGGC TCCGTCGCCG	AGAAGAAGAT TCTTCTTCTA	GCAGGCAGCT CGTCCGTCTA
1801	GAGTTGTTGT CTCAACAACA	ATTCTGATAA TAAGACTATT	GAGTCAGAGG CTCAGTCTCC	TAACTCCCGT ATTGAGGGCA	TGCGGTGCTG ACGCCACGAC

pCMV-NS34A

1851	TTAACGGTGG	AGGGCAGTGT	AGTCTGAGCA	GTACTCGTTG	CTGCCGCGCG
	AATTGCCACC	TCCCGTCACA	TCAGACTCGT	CATGAGCAAC	GACGGCGCGC
<hr/>					
1901	CGCCACCAGA	CATAATAGCT	GACAGACTAA	CAGACTGTTC	CTTTCCATGG
	GCGGTGGTCT	GTATTATCGA	CTGTCTGATT	GTCTGACAAG	GAAAGGTACC
<hr/>					
+2					M A P
			EcoRI		
			-----		
1951	GTCTTTTCTG	CAGTCACCGT	CGTCGACCTA	AGAATTCACC	ATGGCGCCCA
	CAGAAAAGAC	GTCAGTGGCA	GCAGCTGGAT	TCTTAAGTGG	TACCGCGGGT
<hr/>					
+2	I T A Y	A Q Q	T R G L	L G C	I I T
2001	TCACGGCGTA	CGCCAGCAG	ACAAGGGGCC	TCCTAGGGTG	CATAATCACC
	AGTGCCGCAT	GCGGGTCGTC	TGTTCCCCCG	AGGATCCAC	GTATTAGTGG
<hr/>					
+2	S L T G	R D K	N Q V	E G E V	Q I V
2051	AGCCTAACTG	GCCGGGACAA	AAACCAAGTG	GAGGGTGAGG	TCCAGATTGT
	TCCGATTGAC	CGGCCCTGTT	TTTGTTTCAC	CTCCCACTCC	AGGTCTAACA
<hr/>					
+2	S T A	A Q T F	L A T	C I N	G V C
2101	GTCAACTGCT	GCCCAAACCT	TCCTGGCAAC	GTGCATCAAT	GGGGTGTGCT
	CAGTTGACGA	CGGGTTTGGA	AGGACCGTTG	CACGTAGTTA	CCCCACACGA
<hr/>					
+2	W T V Y	H G A	G T R T	I A S	P K G
2151	GGACTGTCTA	CCACGGGGCC	GGAACGAGGA	CCATCGCGTC	ACCCAAGGGT
	CCTGACAGAT	GGTGCCCCGG	CCTTGCTCCT	GGTAGCGCAG	TGGGTTCCCA
<hr/>					
-2	P V I Q	M Y T	N V D	Q D L V	G W P
2201	CCTGTCATCC	AGATGTATAC	CAATGTAGAC	CAAGACCTTG	TGGGCTGGCC
	GGACAGTAGG	TCTACATATG	GTTACATCTG	GTTCTGGAAC	ACCCGACCGG
<hr/>					
+2	A S Q	G T R S	L T P	C T C	G S S
2251	CGCTTCGCAA	GGTACCCGCT	CATTGACACC	CTGCACTTGC	GGCTCCTCGG
	GCGAAGCGTT	CCATGGGCGA	GTAAGTGTGG	GACGTGAACG	CCGAGGAGCC
<hr/>					
+2	D L Y L	V T R	H A D V	I P V	R R R
2301	ACCTTTACCT	GGTCACGAGG	CACGCCGATG	TCATTCCCGT	GCGCCGGCGG
	TGGAAATGGA	CCAGTGCTCC	GTGCGGCTAC	AGTAAGGGCA	CGCGCCGCC
<hr/>					
+2	G D S R	G S L	L S P	R P I S	Y L K
2351	GGTGATAGCA	GGGGCAGCCT	GCTGTCGCCC	CGGCCCATTT	CCTACTTGAA
	CCACTATCGT	CCCCGTCGGA	CGACAGCGGG	GCCGGGTAAA	GGATGAACTT
<hr/>					
+2	G S S	G G P L	L C P	A G H	A V G
2401	AGGCTCCTCG	GGGGGTCCGC	TGTTGTGCCC	CGCGGGGCAC	GCCGTGGGCA
	TCCGAGGAGC	CCCCCAGGCG	ACAACACGGG	GCGCCCCGTG	CGGCACCCGT
<hr/>					
+2	I F R A	A V C	T R G V	A K A	V D F
2451	TATTTAGGGC	CGCGGTGTGC	ACCCGTGGAG	TGGCTAAGGC	GGTGGACTTT
	ATAAATCCCG	GCGCCACACG	TGGGCACCTC	ACCGATTCCG	CCACCTGAAA
<hr/>					
+2	I P V E	N L E	T T M	R S P V	F T D
2501	ATCCCTGTGG	AGAACCTAGA	GACAACCATG	AGGTCCCCCG	TGTTACGGGA
	TAGGGACACC	TCTTGATCT	CTGTTGGTAC	TCCAGGGGCC	ACAAGTGCCT

**pCMV-NS34A**

+2	N G S S	P P V V	P Q S S	F Q V	A H L
2551	TAACCTCTCT ATTGAGGAGA	CCACCAGTAG GGTGGTCATC	TGCCCCAGAG ACGGGGTCTC	CTTCCAGGTG GAAGGTCCAC	GCTCACCTCC CGAGTGGAGG
+2	H A P T	G S G	K S T	K V P A	A Y A
2601	ATGCTCCAC TACGAGGGTG	AGGCAGCGGC TCCGTCGCCG	AAAAGCACCA TTTTCTGGT	AGGTCCCGG TCCAGGGCCG	TGCATATGCA ACGTATACGT
+2	A Q G Y	K V L	V L N	P S V A	A T L
2651	GCTCAGGGCT CGAGTCCCGA	ATAAGGTGCT TATTCCACGA	AGTACTCAAC TCATGAGTTG	CCCTCTGTTG GGGAGACAAC	CTGCAACACT GACGTTGTGA
+2	G F G	A Y M S	K A H	G I D	P N I
2701	GGGCTTTGGT CCCGAAACCA	GCTTACATGT CGAATGTACA	CCAAGGCTCA GGTTCCGAGT	TGGGATCGAT ACCCTAGCTA	CCTAACATCA GGATTGTAGT
+2	R T G V	R T I	T T G S	P I T	Y S T
2751	GGACCGGGGT CCTGSCCCA	GAGAACAATT CTCTTGTTAA	ACCACTGGCA TGGTGACCGT	GCCCCATCAC CGGGTAGTGT	GTACTCCACC CATGAGGTGG
+2	Y G K F	L A D	G G C	S G G A	Y D I
2801	TACGGCAAGT ATGCCGTTCA	TCCTTGCCGA AGGAACGGCT	CGGCGGGTGC GCCGCCACG	TCGGGGGGCG AGCCCCCGC	CTTATGACAT GAATACTGTA
+2	I I C	D E C H	S T D	A T S	I L G
2851	AATAATTTGT TTATTAAACA	GACGAGTGCC CTGCTCACGG	ACTCCACGGA TGAGGTGCCT	TGCCACATCC ACGGTGTAGG	ATCTTGGGCA TAGAACCCGT
+2	I G T V	L D Q	A E T A	G A R	L V V
2901	TTGGCACTGT AACCCTGACA	CCTTGACCAA GGAAGTGGTT	GCAGAGACTG CGTCTCTGAC	CGGGGGCGAG GCCCCCGCTC	ACTGGTTGTG TGACCAACAC
+2	L A T A	T P P	G S V	T V P H	P N I
2951	CTCGCCACCG GAGCGGTGGC	CCACCCCTCC GGTGGGGAGG	GGGCTCCGTC CCCGAGGCAG	ACTGTGCCCC TGACACGGGG	ATCCCAACAT TAGGGTTGTA
+2	E E V	A L S T	T G E	I P F	Y G K
3001	CGAGGAGGTT GCTCCTCCAA	GCTCTGTCCA CGAGACAGGT	CCACCGGAGA GGTGGCCTCT	GATCCCTTTT CTAGGGAAAA	TACGGCAAGG ATGCCGTTCC
+2	A I P L	E V I	K G G R	H L I	F C H
3051	CTATCCCCCT GATAGGGGGA	CGAAGTAATC GCTTCATTAG	AAGGGGGGGA TTCCCCCCT	GACATCTCAT CTGTAGAGTA	CTTCTGTCAT GAAGACAGTA
+2	S K K K	C D E	L A A	K L V A	L G I
3101	TCAAAGAAGA AGTTTCTTCT	AGTGCGACGA TCACGCTGCT	ACTCGCCGCA TGAGCGGCGT	AAGCTGGTCG TTCGACCAGC	CATTGGGCAT GTAACCCGTA
+2	N A V	A Y Y R	G L D	V S V	I P T
3151	CAATGCCGTG GTTACGGCAC	GCCTACTACC CGGATGATGG	GCGGTCTTGA CGCCAGAACT	CGTGTCCGTC GCACAGGCAG	ATCCCCGACCA TAGGGCTGGT
+2	S G D V	V V V	A T D A	L M T	G Y T
3201	GCGGCGATGT CGCCGCTACA	TGTCGTCGTG ACAGCAGCAC	GCAACCGATG CGTTGGCTAC	CCCTCATGAC GGGAGTACTG	CGGCTATACC GCCGATATGG

pCMV-NS34A

```

+2  G D F D S V I D C N T C V T Q T V
3251  GGCGACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT
      CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

+2  D F S L D P T F T I E T I T L P
3301  CGATTTTCAGC CTTGACCCTA CCTTCACCAT TGAGACAATC ACGCTCCCCC
      GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCGAGGGGG

+2  Q D A V S R T Q R R G R T G R G K
3351  AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG
      TTCTACGAGA GAGGGCGTGA GTTGCAGCCC CGTCCTGACC GTCCCCCTTC

+2  P G I Y R F V A P G E R P S G M F
3401  CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCCT CCGGCATGTT
      GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

+2  D S S V L C E C Y D A G C A W Y
3451  CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG
      GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2  E L T P A E T T V R L R A Y M N T
3501  AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC
      TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

+2  P G L P V C Q D H L E F W E G V F
3551  CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT
      GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTAAAACCC TCCCGCAGAA

+2  T G L T H I D A H F L S Q T K Q
      StuI
      ~~~~~
3601  TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCCAG ACAAAGCAGA
      ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTCGTCT

+2  S G E N L P Y L V A Y Q A T V C A
3651  GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT
      CACCCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCGA

+2  R A Q A P P P S W D Q M W K C L I
3701  AGGGGCTCAAG CCCCTCCCCC ATCGTGGGAC CAGATGTGGA AGTGTTTGAT
      TCCCGAGTTC GGGGAGGGGG TAGCACCTTG GTCTACACCT TCACAACTA

+2  R L K P T L H G P T P L L Y R L
3751  TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG
      AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGGACGAT ATGTCTGACC

+2  G A V Q N E I T L T H P V T K Y I
3801  GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC
      CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCAGTG GTTTATGTAG

+2  M T C M S A D L E V V T S T W V L
3851  ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT
      TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

+2  V G G V L A A L A A Y C L S T G
3901  CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCCTG TCAACAGGCT
      GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA
  
```



pCMV-NS34A

	+2	C	V	V	I	V	G	R	V	V	L	S	G	K	P	A	I	I	
3951		GCGTGGTCAT	AGTGGGCAGG	GTCGTCTTGT	CCGGGAAGCC	GGCAATCATA	CGCACCAGTA	TCACCCGTCC	CAGCAGAACA	GGCCCTTCGG	CCGTTAGTAT								
<hr/>																			
	+2	P	D	R	E	V	L	Y	R	E	F	D	E	M	E	E	C		
4001		CCTGACAGGG	AAGTCCTCTA	CCGAGAGTTC	GATGAGATGG	AAGAGTGCTA	GGACTGTCCC	TTCAGGAGAT	GGCTCTCAAG	CTACTCTACC	TTCTCACGAT								
<hr/>																			
		BamHI		MluI															
		-----		-----															
4051		GGATCCACTA	CGCGTTAGAG	CTCGCTGATC	AGCCTCGACT	GTGCCTTCTA	CCTAGGTGAT	GCGCAATCTC	GAGCGACTAG	TCGGAGCTGA	CACGGAAGAT								
<hr/>																			
4101		GTTGCCAGCC	ATCTGTTGTT	TGCCCCCTCCC	CCGTGCCTTC	CTTGACCCCTG	CAACGGTCCG	TAGACAACAA	ACGGGGGAGGG	GGCACGGAAG	GAAGTGGGAC								
<hr/>																			
4151		GAAGGTGCCA	CTCCCACTGT	CCTTTCTCTAA	TAAAATGAGG	AAATTGCATC	CTTCCACGGT	GAGGGTGACA	GGAAAGGATT	ATTTTACTCC	TTTAACGTAG								
<hr/>																			
4201		GCATTGTCTG	AGTAGGTGTC	ATTCTATTCT	GGGGGGTGGG	GTGGGGCAGG	CGTAACAGAC	TCATCCACAG	TAAGATAAGA	CCCCCACC	CACCCCGTCC								
<hr/>																			
4251		ACAGCAAGGG	GGAGGATTGG	GAAGACAATA	GCAGGCATGC	TGGGGAGCTC	TGTCGTTCCC	CCTCCTAACC	CTTCTGTTAT	CGTCCGTACG	ACCCCTCGAG								
<hr/>																			
4301		TTCCGCTTCC	TCGCTCACTG	ACTCGCTGCG	CTCGGTCGTT	CGGCTGCGGC	AAGCGGAAGG	AGCGAGTGAC	TGAGCGACGC	GAGCCAGCAA	GCCGACGCCG								
<hr/>																			
4351		GAGCGGTATC	AGCTCACTCA	AAGGCGGTAA	TACGGTTATC	CACAGAATCA	CTCGCCATAG	TCGAGTGAGT	TTCCGCCATT	ATGCCAATAG	GTGTCTTAGT								
<hr/>																			
4401		GGGGATAACG	CAGGAAAGAA	CATGTGAGCA	AAAGGCCAGC	AAAAGGCCAG	CCCCTATTGC	GTCCTTTCTT	GTACACTCGT	TTTCCGGTGC	TTTTCCGGTC								
<hr/>																			
4451		GAACCGTAAA	AAGGCCGCGT	TGCTGGCGTT	TTTCCATAGG	CTCCGCCCCC	CTTGGCATT	TTCCGGCGCA	ACGACCGCAA	AAAGGTATCC	GAGGCGGGGG								
<hr/>																			
4501		CTGACGAGCA	TCACAAAAAT	CGACGCTCAA	GTCAGAGGTG	GCGAAACCCG	GACTGCTCGT	AGTGTTTTTA	GCTGCGAGTT	CAGTCTCCAC	CGCTTTGGGC								
<hr/>																			
4551		ACAGGACTAT	AAAGATACCA	GGCGTTTCCC	CCTGGAAGCT	CCCTCGTGCG	TGTCCTGATA	TTTCTATGGT	CCGCAAAGGG	GGACCTTCGA	GGGAGCACGC								
<hr/>																			
4601		CTCTCCTGTT	CCGACCCTGC	CGCTTACCGG	ATACCTGTCC	GCCTTTCTCC	GAGAGGACAA	GGCTGGGACG	GCGAATGGCC	TATGGACAGG	CGGAAAGAGG								
<hr/>																			
4651		CTTCGGGAAG	CGTGCGCCTT	TCTCAATGCT	CACGCTGTAG	GTATCTCAGT	GAAGCCCTTC	GCACCGCGAA	AGAGTTACGA	GTGCGACATC	CATAGAGTCA								
<hr/>																			
4701		TCGGTGTAGG	TCGTTCGCTC	CAAGCTGGGC	TGTGTGCACG	AACCCCCCGT	AGCCACATCC	AGCAAGCGAG	GTTGACCCG	ACACACGTGC	TTGGGGGGCA								
<hr/>																			
4751		TCAGCCCGAC	CGCTGCGCCT	TATCCGGTAA	CTATCGTCTT	GAGTCCAACC	AGTCGGGCTG	GCGACGCGGA	ATAGGCCATT	GATAGCAGAA	CTCAGGTTGG								
<hr/>																			
4801		CGGTAAGACA	CGACTTATCG	CCACTGGCAG	CAGCCACTGG	TAACAGGATT	GCCATTCTGT	GCTGAATAGC	GGTGACCGTC	GTCGGTGACC	ATTGTCCTAA								

pCMV-NS34A

4851	AGCAGAGCGA	GGTATGTAGG	CGGTGCTACA	GAGTTCTTGA	AGTGGTGGCC
	TCGTCTCGCT	CCATACATCC	GCCACGATGT	CTCAAGAACT	TCACCACCGG
4901	TAACTACGGC	TACACTAGAA	GGACAGTATT	TGGTATCTGC	GCTCTGCTGA
	ATTGATGCCG	ATGTGATCTT	CCTGTCATAA	ACCATAGACG	CGAGACGACT
4951	AGCCAGTTAC	CTTCGGAAAA	AGAGTTGGTA	GCTCTTGATC	CGGCAAACAA
	TCGGTCAATG	GAAGCCTTTT	TCTCAACCAT	CGAGAAGTAG	GCCGTTTGT
5001	ACCACCGCTG	GTAGCGGTGG	TTTTTTTGT	TGCAAGCAGC	AGATTACGCG
	TGGTGGCGAC	CATCGCCACC	AAAAAAACAA	ACGTTCGTCG	TCTAATGCGC
5051	CAGAAAAAAA	GGATCTCAAG	AAGATCCTTT	GATCTTTTCT	ACGGGGTCTG
	GTCTTTTTTT	CCTAGAGTTC	TTCTAGGAAA	CTAGAAAAGA	TGCCCCAGAC
5101	ACGCTCAGTG	GAACGAAAAC	TCACGTTAAG	GGATTTTGGT	CATGAGATTA
	TGCGAGTCAC	CTTGCTTTTG	AGTGCAATTC	CCTAAAACCA	GTACTCTAAT
5151	TCAAAAAGGA	TCTTCACCTA	GATCCTTTTA	AATTAAAAAT	GAAGTTTAA
	AGTTTTTCCT	AGAAGTGGAT	CTAGGAAAAA	TTAATTTTAA	CTTCAAAATT
5201	ATCAATCTAA	AGTATATATG	AGTAAACTTG	GTCTGACAGT	TACCAATGCT
	TAGTTAGATT	TCATATATAC	TCATTGGAAC	CAGACTGTCA	ATGGTTACGA
5251	TAATCAGTGA	GGCACCTATC	TCAGCGATCT	GTCTATTTTCG	TTCATCCATA
	ATTAGTCACT	CCGTGGATAG	AGTCGCTAGA	CAGATAAAGC	AAGTAGGTAT
5301	GTTGCCTGAC	TCCCCGTCGT	GTAGATAACT	ACGATACGGG	AGGGCTTACC
	CAACGGACTG	AGGGGCAGCA	CATCTATTGA	TGCTATGCCC	TCCCGAATGG
5351	ATCTGGCCCC	AGTGCTGCAA	TGATACCGCG	AGACCCACGC	TCACCGGCTC
	TAGACCGGGG	TCACGACGTT	ACTATGGCGC	TCTGGGTGCG	AGTGGCCGAG
5401	CAGATTTATC	AGCAATAAAC	CAGCCAGCCG	GAAGGGCCGA	GCGCAGAAGT
	GTCTAAATAG	TCGTTATTTG	GTCGGTCGGC	CTTCCCGGCT	CGCGTCTTCA
5451	GGTCCTGCAA	CTTTATCCGC	CTCCATCCAG	TCTATTAATT	GTTGCCGGGA
	CCAGGACGTT	GAAATAGGCG	GAGGTAGGTC	AGATAATTAA	CAACGGCCCT
5501	AGCTAGAGTA	AGTAGTTCGC	CAGTTAATAG	TTTGCGCAAC	GTTGTTGCCA
	TCGATCTCAT	TCATCAAGCG	GTCATTATC	AAACGCGTTG	CAACAACGGT
5551	TTGCTACAGG	CATCGTGGTG	TCACGCTCGT	CGTTTGGTAT	GGCTTCATTC
	AACGATGTCC	GTAGCACCAC	AGTGCGAGCA	GCAAACCATA	CCGAAGTAAG
5601	AGCTCCGGTT	CCCAACGATC	AAGGCGAGTT	ACATGATCCC	CCATGTTGTG
	TCGAGGCCAA	GGGTTGCTAG	TTCCGCTCAA	TGTACTAGGG	GGTACAACAC
5651	CAAAAAGCG	GTTAGCTCCT	TCGGTCCTCC	GATCGTTGTC	AGAAAGTAAGT
	GTTTTTTCGC	CAATCGAGGA	AGCCAGGAGG	CTAGCAACAG	TCTTCATTCA
5701	TGGCCGCAGT	GTTATCACTC	ATGGTTATGG	CAGCACTGCA	TAATTCTCTT
	ACCGGCGTCA	CAATAGTGAG	TACCAATACC	GTCGTGACGT	ATTAAGAGAA
5751	ACTGTCATGC	CATCCGTAAG	ATGCTTTTCT	GTGACTGGTG	AGTACTCAAC
	TGACAGTACG	GTAGGCATTG	TACGAAAAGA	CACTGACCAC	TCATGAGTTG

pCMV-NS34A

5801	CAAGTCATTC	TGAGAATAGT	GTATGCGGCG	ACCGAGTTGC	TCTTGCCCGG
	GTTCAAGTAAG	ACTCTTATCA	CATACGCCGC	TGGCTCAACG	AGAACGGGCG
5851	CGTCAATACG	GGATAATACC	GCGCCACATA	GCAGAACTTT	AAAAGTGCTC
	GCAGTTATGC	CCTATTATGG	CGCGGTGTAT	CGTCTTGAAA	TTTTACAGAG
5901	ATCATTGGAA	AACGTTCTTC	GGGGCGAAAA	CTCTCAAGGA	TCTTACCGCT
	TAGTAACCTT	TTGCAAGAAG	CCCCGCTTTT	GAGAGTTCCT	AGAATGGCGA
5951	GTTGAGATCC	AGTTCGATGT	AACCCACTCG	TGCACCCAAC	TGATCTTCAG
	CAACTCTAGG	TCAAGCTACA	TTGGGTGAGC	ACGTGGGTTG	ACTAGAAGTC
6001	CATCTTTTAC	TTTCACCAGC	GTTTCTGGGT	GAGCAAAAAC	AGGAAGGCAA
	GTAGAAAATG	AAAGTGGTCG	CAAAGACCCA	CTCGTTTTTG	TCCTTCCGTT
6051	AATGCCGCAA	AAAAGGGAAT	AAGGGCGACA	CGGAAATGTT	GAATACTCAT
	TTACGGCGTT	TTTTCCCTTA	TTCCCGCTGT	GCCTTTACAA	CTTATGAGTA
6101	ACTCTTCCTT	TTTCAATATT	ATTGAAGCAT	TTATCAGGGT	TATTGTCTCA
	TGAGAAGGAA	AAAGTTATAA	TAACCTCGTA	AATAGTCCCA	ATAACAGAGT
6151	TGAGCGGATA	CATATTTGAA	TGTATTTAGA	AAAATAAACA	AATAGGGGTT
	ACTCGCCTAT	GTATAAACTT	ACATAAATCT	TTTTATTTGT	TTATCCCCAA
6201	CCGCGCACAT	TTCCCCGAAA	AGTGCCACCT	GACGTCTAAG	AAACCATTAT
	GGCGCGTGTA	AAGGGGCTTT	TCACGGTGGA	CTGCAGATTC	TTTGGTAATA
6251	TATCATGACA	TTAACCTATA	AAAATAGGCG	TATCACGAGG	CCCTTTCGTC
	ATAGTACTGT	AATTGGATAT	TTTTATCCGC	ATAGTGCTCC	GGGAAAGCAG

Diagram 1

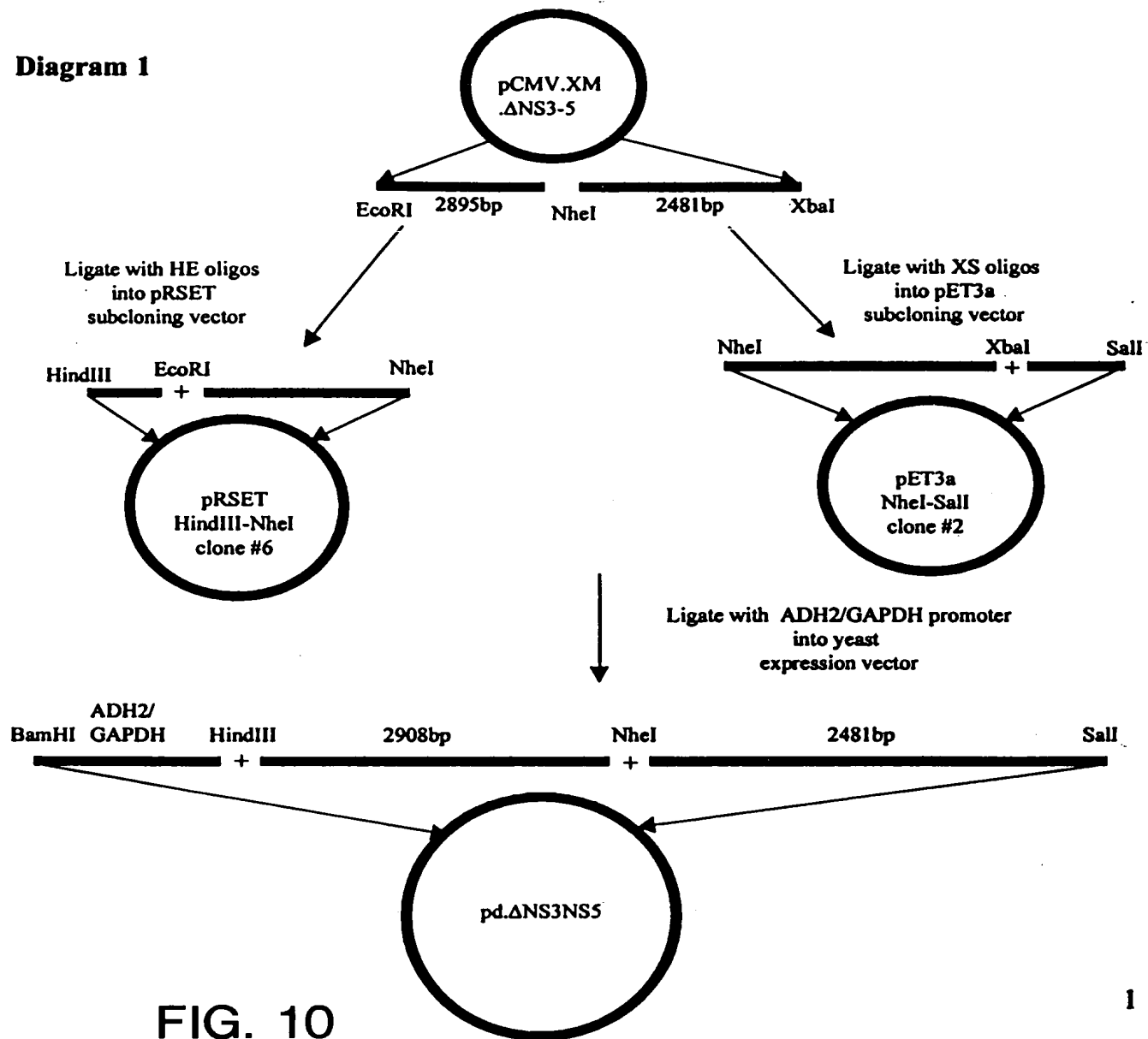


FIG. 10

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal

2 AGCTTACAAAACAAATTCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA  
 TCGAATGTTTTGTTAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCCACGATCAT  
 ^ ^ ^

1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,

LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly

62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG  
 GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC

IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr

122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTAC  
 TAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG  
 ^

122 CLAI,

SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle

182 TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA  
 AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTAT

IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu

242 ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT  
 TAAACACTGCTCACGGTGAGGTGCCTACGGTGTTAGGTAGTAACCCGTAACCGTGACAGGAA

AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly

302 GACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC  
 CTGGTTCGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCC  
 ^

309 ALWN1,

SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle

362 TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATC  
 AGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG

ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe

422 CCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC  
 GGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG

CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn

482 TGTCAATCAAAGAAGAAGTGCAGCAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAAT  
 ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA

AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal

542 GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGCGGATGTTGTC  
 CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAG  
 ^ ^

556 SAC2, 566 DRD1,

ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp

602 GTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGAC  
 CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG  
 ^

621 BSPH1,

CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

662 TGCAATACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAG  
ACGTTTATGCACACAGTGGGTCTGTCTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTC

ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg  
722 ACAATCACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG  
TGTTAGTGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC

GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp  
782 GGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTTCGAC  
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG

822 BGLI, 839 DRD1,

SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla  
842 TCGTCCGTCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCC  
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGG

887 SACI,

GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp  
902 GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC  
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCTG

937 SMAI XMAI,

HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu  
962 CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTA  
GTAGAACTTAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT

991 STUI,

SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal  
1022 TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCCTTACCTGGTAGCGTACCAAGCCACCGTG  
AGGGTCTGTTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTTCGGTGGCAC

1075 DRA3,

CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg  
1082 TGCGCTAGGGCTCAAGCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTTCGC  
ACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCG

LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn  
1142 CTCAAGCCCACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAGAAT  
GAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA

1156 NCOI,

GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu  
1202 GAAATCACCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTG  
CTTTAGTGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAC

1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

GluValValThrS rThrTrpValL uValGlyGlyValLeuAlaAlaLeuAlaAlaTyr  
1262 GAGGTCGTACAGACACCTGGGTGCTCGTTGGCGGCGTCCCTGGCTGCTTTGGCCGCGTAT  
CTCCAGCAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATA

1322 CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysPr Ala  
 TGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCA  
 ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT  
 1375 NAEI,  
 1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln  
 ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAG  
 TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC  
 1391 DRD1,  
 1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu  
 CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC  
 GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAG  
 1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn  
 GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGTGTCCAGACCAAC  
 CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG  
 1508 PSTI, 1513 TTH3I,  
 1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln  
 TGGCAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAA  
 ACCGTTTTTGTAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTT  
 1571 XHOI, 1592 NDEI,  
 1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe  
 TACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT  
 ATGAACCGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACCTACCGAAAA  
 1649 BSTE2,  
 1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly  
 ACAGCTGCTGTCAACAGCCCACTAACCCTAGCCAAACCCCTCCTCTTCAACATATTGGGG  
 TGTCGACGACAGTGGTCCGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCC  
 1683 ALWN1 PVU2,  
 1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly  
 GGGTGGGTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC  
 CCCACCCACCGACGGGTGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCG  
 1800 ESP1,  
 1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla  
 TTAGCTGGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA  
 AATCGACCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT  
 1808 KAS1 NARI,  
 1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal  
 GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC  
 CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG

1884 SACI, 1905 BSPH1,

1922 ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal  
 CCCTCCACGGAGGACCTGGTCAATCTACTGCCCCCATCCTCTCGCCCCGAGCCCTCGTA  
 GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT  
 ^

1934 TTH3I,

1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal  
 GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCCGGCGAGGGGGCAGTG  
 CAGCCGCACCAGACACGTCGTTATGACGCGGGCGGTGCAACCGGGCCCCGCTCCCCCGTCAC  
 ^ ^

2010 NAEI, 2023 SMAI XMAI,

2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis  
 CAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCACGCAC  
 GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTG  
 ^ ^

2073 SMAI XMAI, 2099 DRA3,

2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal  
 TACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTA  
 ATGCACGGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACAT  
 ^

2121 PVU2,

2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer  
 ACCCAGCTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCC  
 TGGGTTCGAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGG  
 ^ ^

2165 ALWN1, 2170 MST2,

2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr  
 GGTTTCCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC  
 CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGG  
 ^

2226 ECON1,

2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg  
 TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGC  
 ACCGATTTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCG  
 ^ ^ ^

2291 ESP1, 2306 PVU2, 2316 BAMHI,

2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla  
 GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT  
 CCCATATTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA

2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg  
 GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGG  
 CTCTAGTGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC  
 ^ ^ ^

2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,

2462 AsnMetTrpSerGlyThrPhePr IleAsnAlaTyrThrThrGlyProCysThrProLeu  
 AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTT  
 TTGTACACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAA



2486 ASE1, 2503 APAI,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle  
CCTGCGCCGAACCTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA  
GGACGCGGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTAT

2559 PSTI,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro  
AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCG  
TCCGTCCACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGG

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe  
TGCCAGGTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTT  
ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu  
GCGCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAA  
CGCGGGGGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer  
TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCCGAACCGGACGTGGCCGTGTTGACGTCC  
ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGG

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly  
ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGA  
TACGAGTGAAGTAGGGAGGGTATATTGTCTGCTCCGCCGGCCGCTTCCAACCGCTCCCT

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla  
TCACCCCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA  
AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGT

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp  
ACTTGCACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG  
TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACC

2972 ESP1, 2975 SACI,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu  
AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG  
TCCGTCCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGAC

3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu  
GACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCGAGAA  
CTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT

3102 BGL2,

IleLeuArgLysSerArgArgPheAlaGlnAlaLeuPr ValTrpAlaArgPr AspTyr  
 3122 ATCCTGCGGAAGTCTCGGAGATTCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTAT  
 TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATA  
 3149 ALWN1, 3170 EAG1 XMA3,  
 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly  
 3182 AACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGC  
 TTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG  
 3223 HGIE2, 3235 NCOI,  
 CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal  
 3242 TGCCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG  
 ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC  
 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly  
 3302 GTCCTCACTGAATCAACCCTATCTACTGCCTTGCCCGAGCTCGCCACCAGAAGCTTTGGC  
 CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCG  
 3338 SACI, 3352 HIND3,  
 SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro  
 3362 AGTCCTCAACTTCGGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCT  
 TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA  
 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly  
 3422 TCTGGCTGCCCCCGGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCGCTGGAGGGG  
 AGACCGACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC  
 3443 EAM11051,  
 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn  
 3482 GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC  
 CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG  
 3490 BAMHI, 3491 BSAB1, 3493 BSPE1,  
 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro  
 3542 GCGGAGGATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCCG  
 CGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC  
 3595 DRA3,  
 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis  
 3602 TGCGCCGCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC  
 ACGCGCGCCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTG  
 3606 SAC2, 3617 ALWN1, 3661 PFLM1,  
 HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr  
 3662 CACAATTTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAGAAAGTCACA  
 GTGTTAAACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGT  
 3687 DRA3,  
 PheAspArgL uGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA  
 AAACGTGTCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGT  
 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro  
 3782 GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCC  
 CGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGG  
 3822 HIND3,  
 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg  
 3842 CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGA  
 GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCGTTTTCTGCAGGCAACGGTACGGTCT  
 3881 AAT2, 3896 BGLI,  
 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro  
 3902 AAGGCCGTAAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA  
 TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT  
 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly  
 3962 ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGT  
 TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCA  
 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet  
 4022 CGTAAGCCAGCTCGTCTCATCGTGTCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATG  
 GCATTGCGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTAC  
 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe  
 4082 GCTTTGTACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTC  
 CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAG  
 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr  
 4142 CAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC  
 GTTATGAGTGGTCCTGTGCGCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGG  
 4166 ECORI,  
 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle  
 4202 CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC  
 GGTTACCCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAG  
 4235 DRD1, 4242 ALWN1,  
 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle  
 4262 CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCGTGGCCATC  
 GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTGCGGCGCACCGGTAG  
 4307 BGLI, 4314 BALI,  
 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn  
 4322 AAGTCCCTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC  
 TTCAGGGAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTG  
 4351 APAI,  
 CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu  
 4382 TGCGGCTATCGCAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTC

ACGCCGATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAG

4442 ThrCysTyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet  
 ACTTGCTACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATG  
 TGAACGATGTAGTTCCGGGCCCGTCCGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTAC  
 ^

4458 SMAI XMAI,

4502 LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla  
 CTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCG  
 GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGC  
 ^ ^

4514 DRD1, 4517 TTH3I,

4562 AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro  
 GCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCC  
 CGCTCGGACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGG

4622 ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla  
 CCACAACCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCC  
 GGTGTTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGG  
 ^

4643 SACI,

4682 HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla  
 CACGACGGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCG  
 GTGCTGCCGCGACCTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGC  
 ^

4737 NRUI,

4742 ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle  
 AGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATC  
 TCTCGACGCACCCTCTGTCTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAG

4802 MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeu  
 ATGTTTGGCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTCTTTAGCGTCCTT  
 TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAA  
 ^^

4812 PFLM1, 4813 DRA3,

4862 IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer  
 ATAGCCAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCC  
 TATCGGTCCCTGGTCAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGG  
 ^

4899 BGL2,

4922 IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer  
 ATAGAACCCTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCA  
 TATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGT  
 ^

4960 NCOI,

4982 LeuHisSerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly  
 CTCCACAGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGG  
 GAGGTGTCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCC  
 ^ ^

5021 SPHI, 5041 KPNI,

5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla  
 GTACCGCCCTTGGCAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCC  
 CRTGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG  
 5070 APAI, 5097 BALI,

5102 ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys  
 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG  
 TCTCCTCCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTC  
 5119 NDEI,

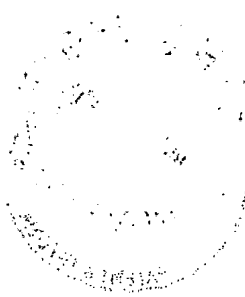
5162 LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla  
 CTCAAACCTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCT  
 GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGA  
 5180 NOTI, 5181 EAGI XMA3, 5188 BALI, 5192 PVU2,

5222 GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp  
 GGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGG  
 CCGATGTCGCCCCCTCTGTAAATAGTGTGCGCACAGAGTACGGGCCGGGGCGACCTAGACC  
 5246 DRA3,

5282 PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
 TTTTGCTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAAGG  
 AAAACGGATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCC  
 5301 PSTI, 5331 HGIE2,

5342 TTGGGGTAAACACTCCGGCCTAAAAAATCTAGAACCCGAGTCGAC  
 AACCCCATTTGTGAGGCCGGATTTTTTTTTTTTAGATCTTGGGCTCAGCTG  
 5378 XBAI, 5390 SALI,

FIG. 11-Page 9



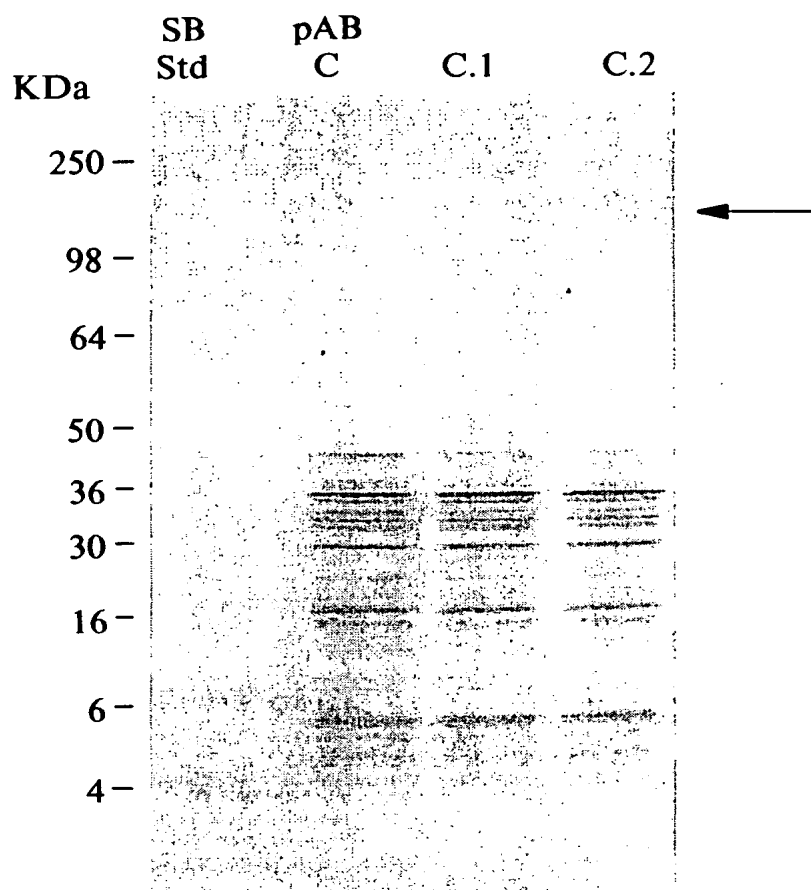


FIG. 12

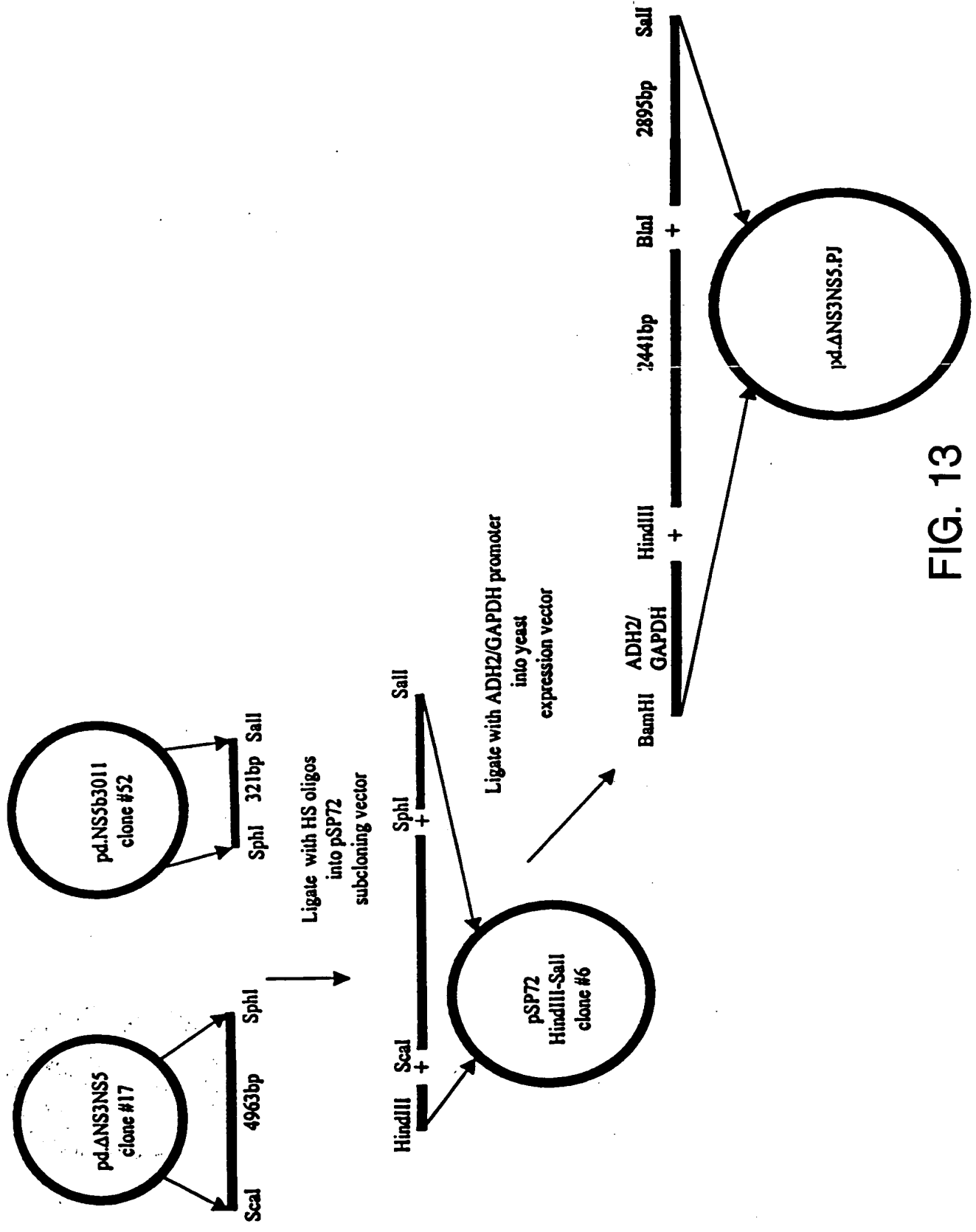


FIG. 13

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,

ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,

ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCAGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA

AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCTAACCCTGACAGGAACCTGGTT

AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,

ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA

TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
^ ^  
550 SAC2, 560 DRD1,

AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,

ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle



662 ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTCTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCAAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACCCCCTGCTATACAGACTGGGCGCTGTTTCAAGTAAATC  
GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTCACTGGTTTATGTAGTACTGTACGTACGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGGCTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA  
 1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 GCTGTACACGACCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
 1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GCGCGCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGACGGGTAT  
 CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
 1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
 ^  
 1928 TTH3I,  
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGTCCCCCGTCACGTCACC  
 ^ ^  
 2004 NAEI, 2017 SMAI XMAI,  
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCACGCACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
 ^ ^  
 2067 SMAI XMAI, 2093 DRA3,  
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 2102 CCGGAGAGCGATGCAGCTGCCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTG  
 ^ ^  
 2115 PVU2, 2159 ALWN1,  
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^ ^  
 2164 MST2, 2220 ECON1,  
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTGTGTCTGCCAGCGCGGGTAT  
 TTTGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA  
 ^ ^ ^  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,  
 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGAC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProL uProAla  
 2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCCTGGC  
 ACCTCACCCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
 ^ ^  
 2480 ASE1, 2497 APAI,  
 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
2642 GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
2702 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
2762 GTAGGGTTCGAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAACATGCAGGTACGAG  
^

2809 AAT2,

ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
2822 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCCTAGTGGG  
^

2850 EAG1 XMA3,

ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
2882 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^

2889 BALI, 2903 NHEI,

ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
2942 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
3002 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTGTTTCACCACTAAGACCTGAGG

PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
3062 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
3122 CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCCCGTTTGGGCGCGGCCGACTATAACCCC

GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
 3143 ALWN1, 3164 EAG1 XMA3,

ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
 GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC

3217 HGIE2, 3229 NCOI,

LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 3242 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 3302 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG

3332 SACI, 3346 HIND3,

SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCTGGAGGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA

3437 EAM11051,

GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
 CCCCTAGGCCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
 3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
 CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
 3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
 AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTCCGTCTTCTTTCAGTGTAACCTG

3681 DRA3,

ArgLeuGlnValLeuAspS rHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
 3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTAAGGAGGTTAAAGCAGCGGGC  
 TCTGACGTTCAAGACCTGTCGGTAATGGTCTGCATGAGTTCCTCCAATTTTCGTCGCCGC

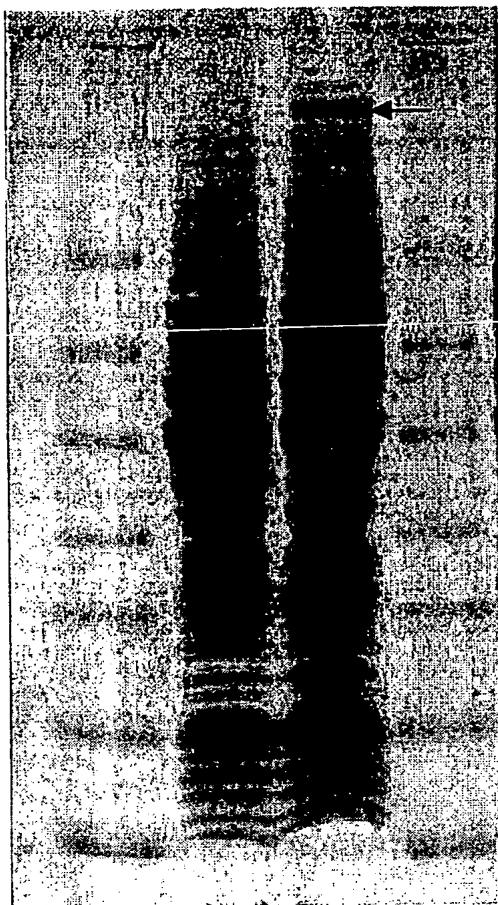
3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
 TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
 AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG  
 3816 HIND3,  
 3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
 TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
 AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
 3875 AAT2, 3890 BGLI,  
 3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
 CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
 3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
 ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTTCGTAAG  
 TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC  
 4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
 CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
 GGTGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC  
 4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
 TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
 ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG  
 4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
 TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
 AGTGGTCCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
 4160 ECORI,  
 4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
 GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
 CCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
 4229 DRD1, 4236 ALWN1,  
 4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC  
 CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG  
 4301 BGLI, 4308 BALI,  
 4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
 CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAGTGCAGG  
 GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTCCCCCTCTTGACGCCG  
 4345 APAI,  
 4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
 TATCGCAGGTGCCGCGGAGCGGCGTACTGACAAGTACTGTGGTAACACCCCTCACTGCG  
 ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
 TACATCAAGGCCCCGGGCAGCCTGTCGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
 ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCCGAGGTCTTGACGTGGTACGAGCAC  
 ^  
 4452 SMAI XMAI,  
  
 4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
 ^ ^  
 4508 DRD1, 4511 TTH3I,  
  
 4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT  
  
 4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCACGAC  
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTCTG  
 ^  
 4637 SACI,  
  
 4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
 GCGGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
 CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
 ^  
 4731 NRUI,  
  
 4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
 CGCACCTCTGTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
  
 4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
 CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
 ^^  
 4806 PFLM1, 4807 DRA3,  
  
 4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
 TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
  
 4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
 ^  
 4954 NCOI,  
  
 4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
 TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTGAACCCCATGGC  
 ^ ^  
 5015 SPHI, 5035 KPNI,  
  
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

5042 CCCTTGGCAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
 GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
 ^ ^  
 5064 APAI, 5091 BALI,  
  
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
 5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
 CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
 ^  
 5113 NDEI,  
  
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
 5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
 GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
 ^ ^ ^  
 5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,  
  
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
 5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
 TCGCCCCCTCTGTAAATAGTGTGCGCACAGAGTACGGGGCCGGGGCGACCTAGACCAAACG  
 ^  
 5240 DRA3,  
  
 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
 5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAATAGTCGAC  
 GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTACTTATCAGCTG  
 ^ ^  
 5295 PSTI, 5336 SALI,

FIG. 14-Page 9





**FIG. 15**

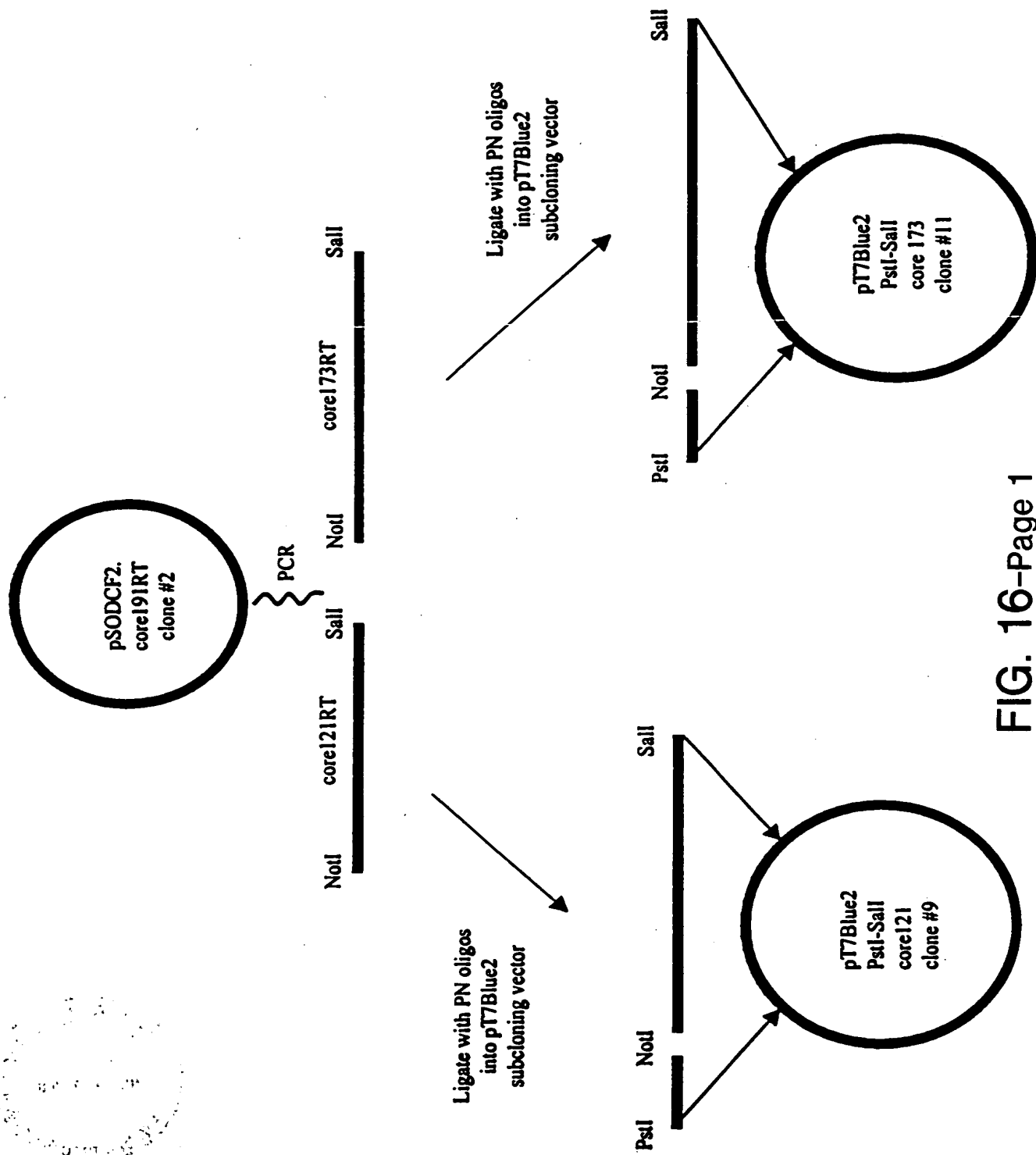
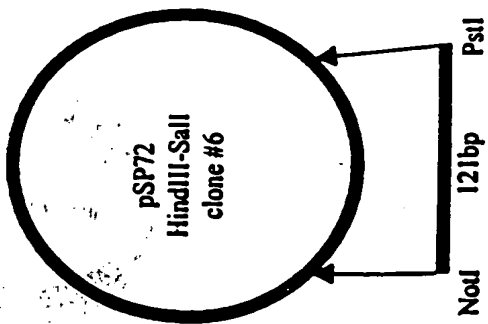


FIG. 16-Page 1



Ligate fragments into pd.ΔNS3NS5.PJ  
NotI-Sall cloning vector.

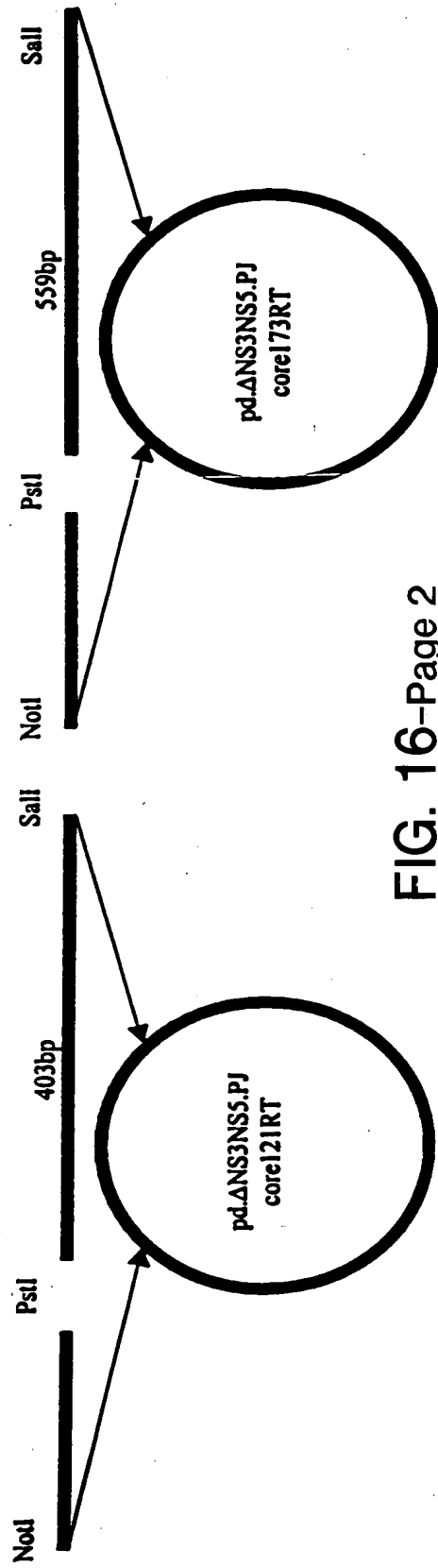


FIG. 16-Page 2

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn

2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
 TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
 ^ ^

1 HIND3, 24 NDEI, 52 SCAI,

ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp

62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
 ^

116 CLAI,

ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr

122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys

182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
 ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGCGAATACTGTATTATTAAACA

AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln

242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCGTAACCGTGACAGGAAGTGGTT

AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal

302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
 CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
 ^

303 ALWN1,

ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe

362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA

TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis

422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 17-Page 1

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 TCAAAGAAGAAGTGGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGGCATCAATGCCGTG  
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
 542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 GCCTACTACCGCGGTCTTGACGTGTCCGTCAATCCCGACCAGCGGCGATGTTGTCTGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,  
  
 602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,  
  
 662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG  
  
 722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGGAAG  
 TGGGAGGGGGTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC  
  
 782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 ^ ^  
 816 BGLI, 833 DRD1,  
  
 842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
 ^  
 881 SACI,  
  
 902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA  
 ^  
 931 SMAI XMAI,  
  
 962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCCTTTCTATCCCAG  
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
 ^  
 985 STUI,  
  
 1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 ^  
 1069 DRA3,  
  
 1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC  
 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTGAGAATGAAATC  
 GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^  
 1150 NCOI,  
 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,  
 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC  
 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 ^  
 1369 NAEI,  
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTGTGAAT  
 ^  
 1385 DRD1,  
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG  
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 ^ ^  
 1502 PSTI, 1507 TTH3I,  
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCCTATGTTATGAAC  
 ^ ^  
 1565 XHOI, 1586 NDEI,  
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 1622 GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA  
 ^ ^  
 1643 BSTE2, 1677 ALWN1 PVU2,  
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 1682 GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCAGCTCGCCGCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTCGAGCGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
 1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGAGGGTAT  
 CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
 1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGCGCGGGCGTGCGGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCACTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
 1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCCGGGCGAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
 2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCACGCACTACGTG  
 TACTTGCGCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
 2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
 2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
 TTTGATTTCGAGTACGGTGTGACGGACCCCTAGGGGAAACACAGGACGGTCGCGCCCATA  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGAC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG  
 ACCTCACCTGGAAGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC  
 ^ ^  
 2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCGCTC  
 ^  
 2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
 CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
 ^  
 2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCATCGCCCGAATTTTTTACAGAATTGGACGGGGTGCGCCTACATAGTTTGCGCCC  
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTACAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
 ^  
 2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCAGGACGTGGCCGTGTTGACGTCCATGCTC  
 CATCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAACAGCAGGTACGAG  
 ^  
 2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGCGGCGAAGGTTGGCGAGGGGATCACCC  
 TGA TAGGGAGGGTATATTGTCGTCTCCGCCGCGCCGCTTCCAACCGCTCCCCTAGTGGG  
 ^  
 2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProS rLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
 ^ ^  
 2889 BALI, 2903 NHEI,



2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
 2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCTCCTCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
 3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC  
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
 3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
 GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
 3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
 3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
 3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
 CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCAGCGCG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGCATGAGTTCCTCCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG  
^ ^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
^

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCTGCGCCGCTCG  
^ ^

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGCTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePhSerValLeuIleAla  
GCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
^

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCATGGC  
^ ^

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^ ^

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCCTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACC GAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA  
^

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGCGTCTCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGTGAGTCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
 ^ ^ ^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
 ATCCCCAAGGCTCGTCCGCCCCGAGGGCAGGACCTGGGCTCAGCCCCGGGTACCCTTGGCCC  
 TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTTGACCCGAGTCGGGCCCCATGGGAACCGGG  
 ^ ^ ^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
 GAGATAACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM  
 CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAAGTAATAGTCG  
 GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTATTATCAGC  
 ^ ^ ^

5650 APAI, 5698 SALI,

5702 AC  
 TG

FIG. 17-Page 10

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn

2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
 TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
 ^ ^ ^  
 1 HIND3, 24 NDEI, 52 SCAI,

ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp

62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
 ^

116 CLAI,

ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr

122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys

182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGCGCTTATGACATAATAATTTGT  
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCAATACTGTATTATTAAACA

AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln

242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACGGTT

AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal

302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
 CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
 ^

303 ALWN1,

ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe

362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA

TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis

422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal

482 TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal

542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^

550 SAC2, 560 DRD1,

AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspS rValIleAspCysAsn

602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^

615 BSPH1,

662 ThrCysValThrGlnThrValAspPheS rLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
 TGCAGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 ^ ^

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCCGGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGGCGGCTCTGA  
 ^

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTGGTAGAA  
 ^

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTTGGGAGGGCGTCTTTACAGGCCCTCACTCATATAGATGCCCACTTTCTATCCAG  
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
 ^

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTCGTCTCACCCCTCTTGGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 ^

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCCTCAAG  
 TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG

CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC  
 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 1369 NAEI,  
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 1385 DRD1,  
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCCTC  
 GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG  
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 1502 PSTI, 1507 TTH3I,  
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 1565 XHOI, 1586 NDEI,  
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACATTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCTGA  
 1643 BSTE2, 1677 ALWN1 PVU2,  
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 1682 GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC  
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 1742 GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTGAGCGGGCGGGGGCCACGGCGATGACGGAACACCCGCGACCGAATCGA  
 1794 ESP1,  
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 1802 GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCTCATAGACATCCTTGCAGGGTAT  
 CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATATA  
 1802 KAS1 NARI,  
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 1878 SACI, 1899 BSPH1,



1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
 ^  
 1928 TTH3I,  
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
 ^ ^  
 2004 NAEI, 2017 SMAI XMAI,  
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
 ^ ^  
 2067 SMAI XMAI, 2093 DRA3,  
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
 ^ ^  
 2115 PVU2, 2159 ALWN1,  
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^ ^  
 2164 MST2, 2220 ECON1,  
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
 TTTGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT  
 ^ ^ ^  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,  
 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCCTGCG  
 ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGGCCGGGGACATGGGGGGAAGGACGC  
 ^ ^  
 2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAAC TACACG TTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCCGAATTTTTTACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACC GGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAAC TGCAGGTACGAG

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGGCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTGCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleS rValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTTAGAGGCATGGGCGTCTTTAGGAC

3096 BGL2,

ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

3122 CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCCCGTTTGGGCGCGGCCGGACTATAACCCC  
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
 3143 ALWN1, 3164 EAG1 XMA3,  
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 3182 CCGCTAGTGGAGACGTGGAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
 GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
 3217 HGIE2, 3229 NCOI,  
 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 3242 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 3302 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
 3332 SACI, 3346 HIND3,  
 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
 3437 EAM11051,  
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG  
 3589 DRA3, 3600 SAC2,  
 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
 3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
 CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTGGTGTTA  
 3611 ALWN1, 3655 PFLM1,  
 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
 3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTCACATTTGAC  
 AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGT  
 3681 DRA3,  
 ArgLeuGlnValL uAspS rHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
 3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCG

TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG  
^

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
^

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGTTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTCT

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGCCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTGCCCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
^

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCACTGACTCTCGCTGTAGGCATGC  
^

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG  
^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAGTGCAGG  
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValL uThrThrS rCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAAGTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
 TACATCAAGGCCCGGGCAGCCTGTCTGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG  
 ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
 ^  
 4452 SMAI XMAI,  
  
 4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGGCGAGC  
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
 ^ ^  
 4508 DRD1, 4511 TTH3I,  
  
 4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTT  
  
 4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
 ^  
 4637 SACI,  
  
 4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
 GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
 CCGCGACCTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
 ^  
 4731 NRUI,  
  
 4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTGGCTAGGCAACATAATCATGTTT  
 CGCACCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
  
 4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
 CGGGGGTGAGACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
 ^ ^  
 4806 PFLM1, 4807 DRA3,  
  
 4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA  
 TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
  
 4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
 ^  
 4954 NCOI,  
  
 4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAPAACTTGGGGTACCG  
 TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
 ^ ^  
 5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
 CCCTTGGCAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
 GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
 5064 APAI, 5091 BALI,  
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
 5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
 CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
 5113 NDEI,  
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
 5162 CTCACTCCAATAGCGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
 GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
 5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,  
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
 5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
 TCGCCCCCTCTGTAAATAGTGTGCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
 5240 DRA3,  
 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
 5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
 GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
 5295 PSTI,  
 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
 5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
 GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGCTCCTGCAGTTCAAG  
 5380 NOTI, 5381 EAGI XMA3, 5390 AAT2, 5401 SMAI XMAI,  
 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
 5402 CCGGGTGGCGGTTCAGATCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG  
 GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCGGGGATCTAAC  
 5449 APAI,  
 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
 5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTTCGCAACCTCGAGGTAGACGTCAGCCT  
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
 5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,  
 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
 5522 ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
 TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTTGGACCCGAGTCGGGCCCCATGGGAACCGGG  
 5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,  
 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
 5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
 GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
 CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
 GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
 ^ ^  
 5650 APAI, 5696 CLAI,  
 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
 5702 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCTCTT  
 TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA  
 ^ ^  
 5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,  
 GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr  
 5762 GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACAT  
 CCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATA  
 ^ ^  
 5772 BSTXI, 5775 APAI,  
 AlaThrGlyAsnLeuProGlyCysSerOC AM  
 5822 GCAACAGGGAACCTTCCTGGTTGCTCTTAATAGTCGAC  
 CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG  
 ^  
 5854 SALI,

FIG. 18-Page 10

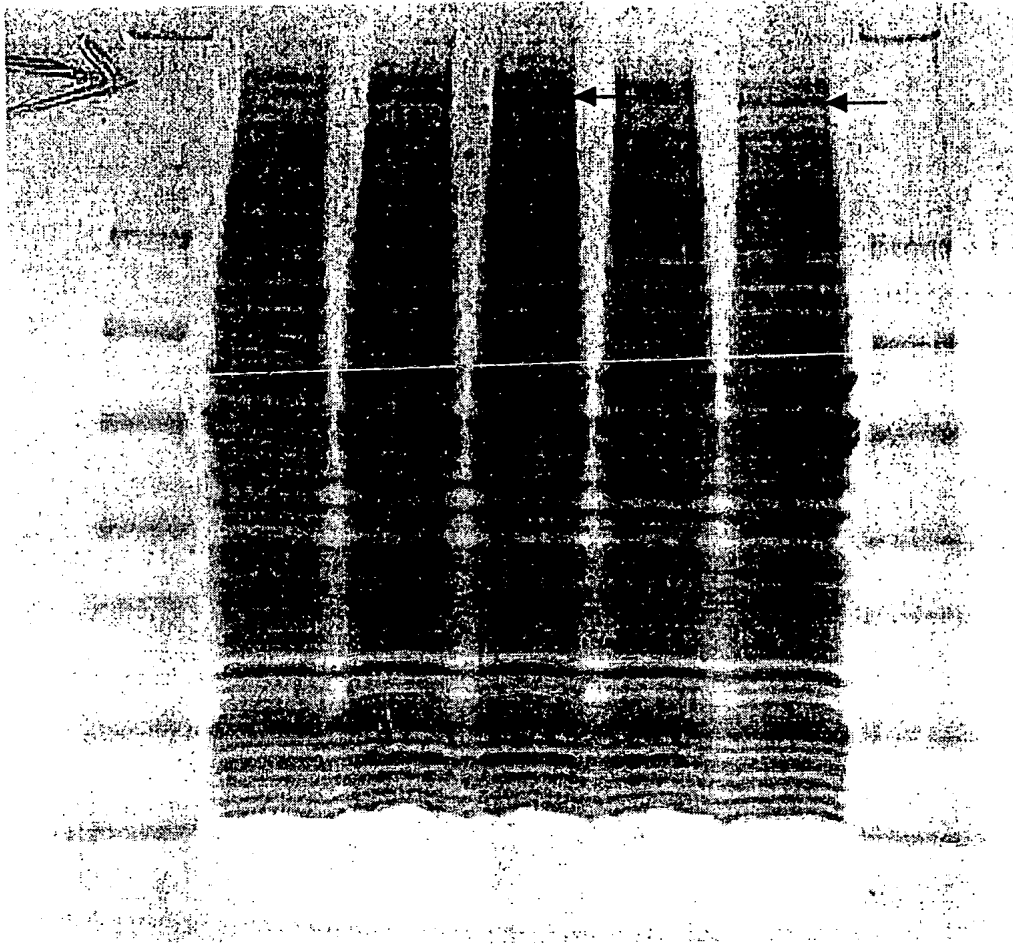


FIG. 19



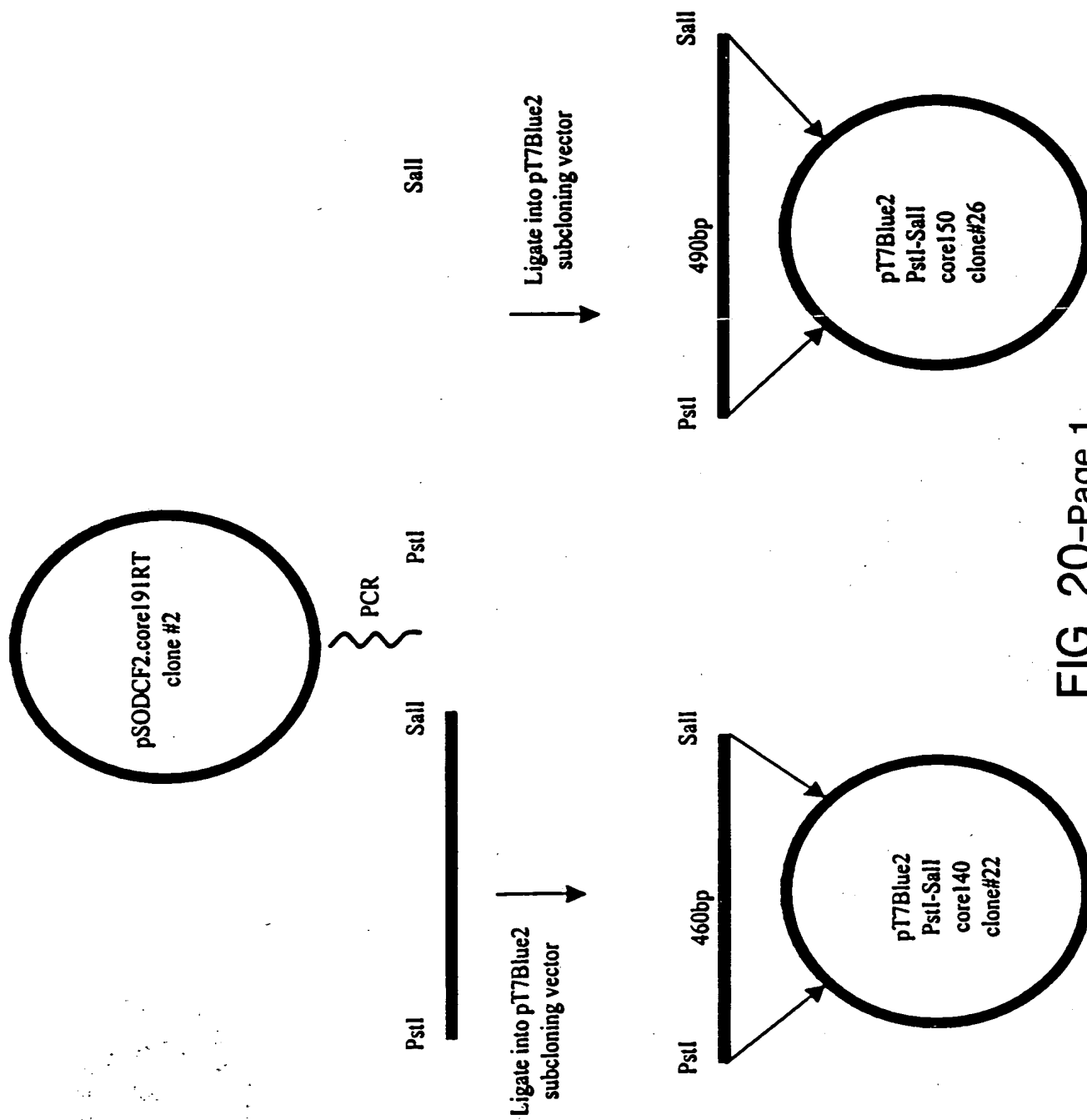
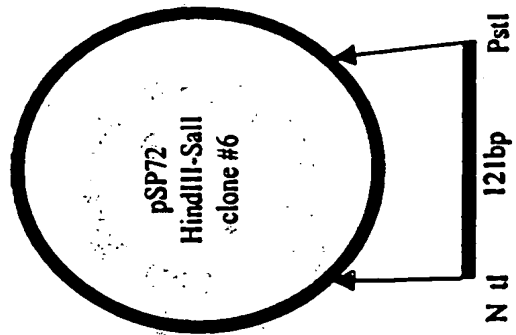


FIG. 20-Page 1



Ligate fragments into pd.NS3NS5.PJ  
NotI-Sall cloning vector.

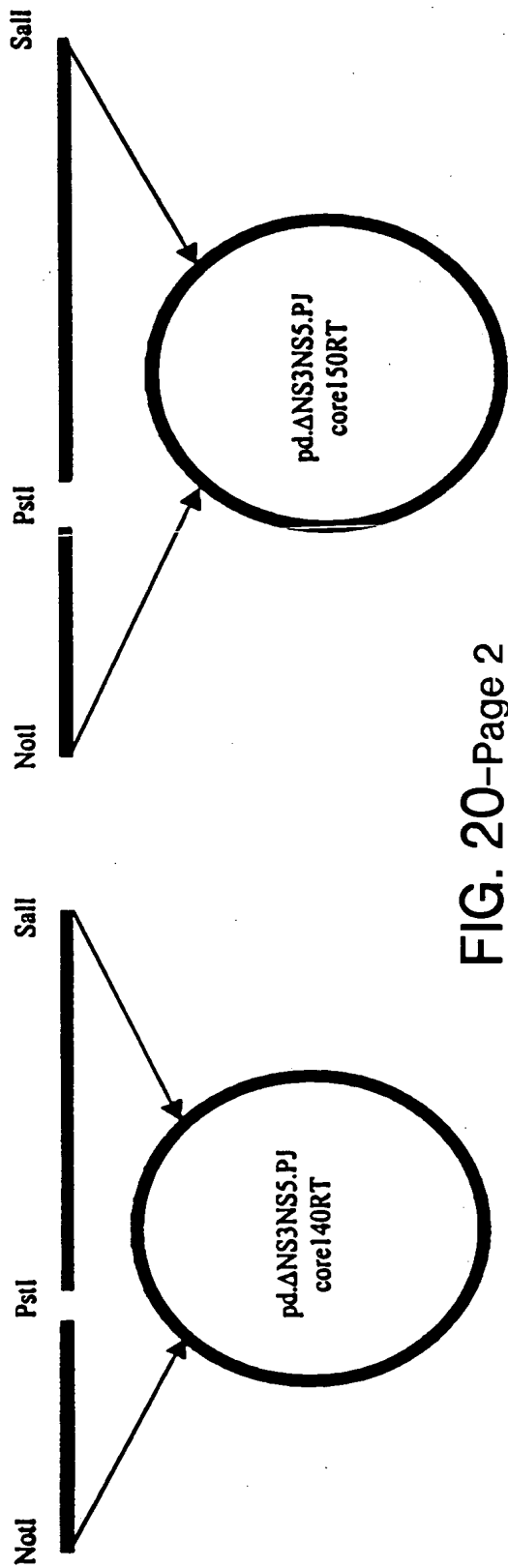


FIG. 20-Page 2

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCTAACCCTGACAGGAACCTGGTT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 21-Page 1

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 TCAAAGAAGAAGTGCAGCAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTCTGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATGAGACAATC  
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
 TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTTGTGGCACCAGGGGGAGCGCCCCCTCCGGCATGTTGCACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 ^ ^  
 816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
 ^  
 881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA  
 ^  
 931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGTC  
 ^  
 985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 ^  
 1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAGAATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^  
 1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGACGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTTCATAGTGGGCAGGGTCGTCTTGTCCGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 ^  
 1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT  
 ^  
 1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCGAAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 ^ ^  
 1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 ^ ^  
 1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA  
 ^ ^  
 1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrS rGlnThrLeuL uPheAsnIleLeuGlyGlyTrp  
 GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTGAGCGGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
 ^  
 1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GCGCGCCGCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
 CCGCGGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
 ^  
 1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GCGCGGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 ^ ^  
 1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
 ^  
 1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCCGGGCGAGGGGGCAGTGCACTGG  
 CACCAGACACGTCGTTATGACGCGGCGGTGCAACCGGGCCGCTCCCCCGTCACGTCACC  
 ^ ^  
 2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCAGCACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
 ^ ^  
 2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
 ^ ^  
 2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^ ^  
 2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
 ^ ^ ^

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
 TTTGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCAT  
 ^ ^ ^  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIle MetHisThrArgCysHisCysGlyAlaGluIle  
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGAC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG  
 ACCTCACCCCTGGAAGGGTAATTACGGATGTGGTGCCCGGGACATGGGGGAAGGACGC  
 ^ ^  
 2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
 ^  
 2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
 CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
 ^  
 2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
 CAGGGTAGCGGGCTTAAAAAGTGCTTAACTGCCCCACGCGGATGTATCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTACAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
 ^  
 2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
 CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGGTACGAG  
 ^  
 2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC  
 TGA TAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG  
 ^  
 2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
 ^ ^  
 2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
 2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCCGCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
 3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC  
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
 3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG  
 GCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
 3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
 3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCCTTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
 3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
 CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrS rTrpThrGlyAlaLeuValThrProCysAla  
 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG



3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGCTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTCT

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGTCGTCTCATCGTGTCCCCGATCTGGGCGTGC GCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG

LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCGGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC

CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCAGGTCTCTCTGCGCCGCTCG

LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGGGGTGT

ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTGTCACAGTCAGCGGGTGCTG

GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACCTGGGATGTTGGGGGGAGCGCTCTCGA

AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA

**AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla**  
**GGCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC**  
**CGGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG**

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCCTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACC GAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCCGCGCAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGCGCTCCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGTGAGTCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
 ^ ^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
 ATCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
 TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
 ^ ^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
 GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
 GGATCGACCCCGGGGTGTCTGGGGGCCGATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
 ^ ^

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM  
 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC  
 TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG  
 ^ ^

5724 HGIE2, 5755 SALI,

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn

2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
 TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
 ^ ^ ^  
 1 HIND3, 24 NDEI, 52 SCAI,

ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp

62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
 ^  
 116 CLAI,

ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr

122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys

182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA

AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln

242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCGTAACCGTGACAGGAAGTGGTT

AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal

302 GCAGAGACTGCGGGGCGGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
 CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
 ^  
 303 ALWN1,

ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe

362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA

TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis

422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 22-Page 1

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
 542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCTCGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,  
  
 602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,  
  
 662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG  
  
 722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
 TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC  
  
 782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCCTCCGGCATGTTTCGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 ^ ^  
 816 BGLI, 833 DRD1,  
  
 842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCAGCCCCGCCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
 ^  
 881 SACI,  
  
 902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA  
 ^  
 931 SMAI XMAI,  
  
 962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
 CTTAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
 ^  
 985 STUI,  
  
 1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTTCGTCTACCCCTCTTGGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 ^  
 1069 DRA3,  
  
 1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTGCGCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProL uLeuTyrArgLeuGlyAlaValGlnAsnGluIl  
 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC  
 GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 ^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 ^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGTCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 ^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 ^ ^

1565 KHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA  
 ^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPh AsnIleLeuGlyGlyTrp  
 GCTGTACACGCCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAACACCCGCGACCGAATCGA  
 ^  
 1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GGC GCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTG CAGGGTAT  
 CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
 ^  
 1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGC GCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 ^ ^  
 1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
 ^  
 1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGGCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
 ^ ^  
 2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
 ^ ^  
 2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
 ^ ^  
 2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^ ^  
 2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuM tProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
 TTTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCAT  
 ^ ^ ^  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,





2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
 ^ ^  
 2966 ESP1, 2969 SACI,  
  
 3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
  
 3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
 ^  
 3096 BGL2,  
  
 3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 CGGAAGTCTCGGAGATTTCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC  
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
 ^ ^  
 3143 ALWN1, 3164 EAG1 XMA3,  
  
 3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACACCTGTGGTCCATGGCTGCCCG  
 GGGCATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
 ^ ^  
 3217 HGIE2, 3229 NCOI,  
  
 3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
  
 3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
 ^ ^  
 3332 SACI, 3346 HIND3,  
  
 3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
  
 3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
 ^  
 3437 EAM11051,  
  
 3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACCGGGAG  
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
 ^ ^ ^  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
  
 3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGCTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTtaggtTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTGCGCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGG  
^ ^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCAGCGGCGTACTGACAAGTCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCCGTGCGACAGCTCGGCGTCCCGAGGTCTTGACGTGGTACGAGCAC  
^

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTGCGCCGCTCG  
^ ^

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTCTG  
^

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGCTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaL uAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAAGTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPN1,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCCGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCCGCGCAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCGCGCTCCTGCAGTTCAAG  
^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
 ^ ^ ^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
 ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
 TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
 ^ ^ ^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
 GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
 GGATCGACCCCGGGGTGTCTGGGGGCCGATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
 ^ ^ ^ ^ ^

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCTCTT  
 TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA  
 ^ ^ ^ ^ ^

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaOC AM  
 GGAGGCGCTGCCAGGGCCTAATAGTCGAC  
 CCTCCGCGACGGTCCCGGATTATCAGCTG  
 ^

5785 SALI,

BEST AVAILABLE COPY

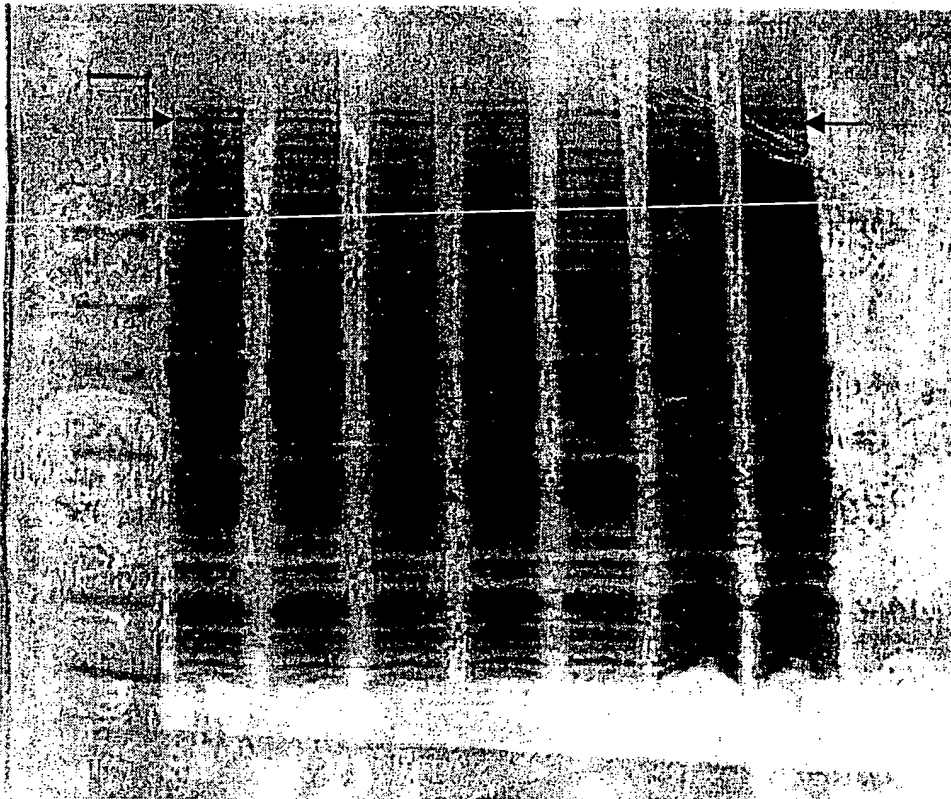


FIG. 23